Structure and Function of Proteins Investigated By Crystallographic and Spectroscopic Time-Resolved Methods

Namrta Purwar
University of Wisconsin-Milwaukee
STRUCTURE AND FUNCTION OF PROTEINS
INVESTIGATED BY CRYSTALLOGRAPHIC AND
SPECTROSCOPIC TIME-RESOLVED METHODS

by

Namrta Purwar

A Dissertation Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy

in Physics

at

The University of Wisconsin-Milwaukee

December 2013
ABSTRACT

STRUCTURE AND FUNCTION OF PROTEINS
INVESTIGATED BY CRYSTALLOGRAPHIC AND
SPECTROSCOPIC TIME-RESOLVED METHODS

by

Namrta Purwar

The University of Wisconsin-Milwaukee, 2013
Under the Supervision of Dr. Marius Schmidt

Biomolecules play an essential role in performing the necessary functions for life. The goal of this thesis is to contribute to an understanding of how biological systems work on the molecular level. We used two biological systems, beef liver catalase (BLC) and photoactive yellow protein (PYP). BLC is a metalloprotein that protects living cells from the harmful effects of reactive oxygen species by converting H₂O₂ into water and oxygen. By binding nitric oxide (NO) to the catalase, a complex was generated that mimics the Cat-H₂O₂ adduct, a crucial intermediate in the reaction promoted by the catalase. The Cat-NO complex is obtained by using a convenient NO generator (1-(N,N-diethylamino)diazen-1-ium-1,2-diolate). Concentrations up to 100~200 mM are reached by using a specially designed glass cavity. With this glass apparatus and DEANO, sufficient NO occupation is achieved and structure determination of the catalase with NO bound to the heme iron becomes possible. Structural changes upon NO binding are
minute. NO has a slightly bent geometry with respect to the heme normal, which results in a substantial overlap of the NO orbitals with the iron-porphyrin molecular orbitals. From the structure of the iron-NO complex, conclusions on the electronic properties of the heme iron can be drawn that ultimately lead to an insight into the catalytic properties of this enzyme.

Enzyme kinetics is affected by additional parameters such as temperature and pH. Additionally, in crystallography, the absorbed X-ray dose may impair protein function. To address the effect of these parameters, we performed time-resolved crystallographic experiments on a model system, PYP. By collecting multiple time-series on PYP at increasing X-ray dose levels, we determined a kinetic dose limit up to which kinetically meaningful X-ray data sets can be collected. From this, we conclude that comprehensive time-series spanning up to 12 orders of magnitude in time can be collected from a single PYP crystal. Time-resolved X-ray data collected at pH’s of 4, 7 and 9 demonstrate that pH alters the kinetics of the PYP photocycle dramatically. At pH 4 the photocycle lasts almost one order of magnitude longer in time compared to pH 7. The final intermediate that accumulates at both pH 7 and pH 4 is absent at pH 9. Results from the dose- and the pH-dependent time-resolved crystallographic experiments show that it is imperative to carefully control the conditions under which time-resolved data are collected. With these considerations we collected a comprehensive time-series from nanoseconds to seconds at 14 different temperature settings from -40 °C to 70 °C. Results from time-resolved crystallography are corroborated by employing time-resolved absorption spectroscopy. For this, absorption spectra on crystals and solution are collected by a fast micro-
spectrophotometer custom-designed in our lab. We identify kinetic phases of the PYP photocycle at all 14 temperature settings. Relaxation times associated with these phases are temperature-dependent and can be fit by the Van’t Hoff-Arrhenius equation. Kinetic modeling yields entropy and enthalpy values at the barriers of the activation solely from the time-resolved crystallographic data. With this, we advance crystallography to a new frontier: the determination of free energy surfaces.

Investigating enzymatic reactions can be challenging, because they are non-cyclic. After one turnover product must be washed away and substrate must be reloaded. A promising approach for routine application can be envisioned at the new 4th generation X-ray sources, such as X-ray free electron lasers (XFELs). With our results we set the scene to comprehensively investigate all kinds of enzymatic reactions with these instruments.
# TABLE OF CONTENTS

ABSTRACT .............................................................................................................................. II

TABLE OF CONTENTS ........................................................................................................... V

LIST OF FIGURES ................................................................................................................ IX

LIST OF TABLES .................................................................................................................... XII

ACKNOWLEDGEMENTS .......................................................................................................... XIII

1 INTRODUCTION .................................................................................................................. 1

1.1 Enzymes: structure and function ................................................................................. 1

1.2 Short history of crystallography .................................................................................. 2

1.3 From static structure to function ................................................................................ 4

1.4 Catalase .......................................................................................................................... 5

1.5 PYP .................................................................................................................................. 9

1.6 Effect of X-rays doses on PYP ..................................................................................... 14

1.7 Effect of pH on the PYP ............................................................................................... 15

1.8 Effect of temperature on PYP ...................................................................................... 16

2 MATERIAL AND METHODS ............................................................................................. 23

2.1 Monochromatic X-ray crystallography ....................................................................... 23

2.1.1 Scattering from an atom ......................................................................................... 23

2.1.2 Scattering from a molecule .................................................................................... 23

2.1.3 Scattering from a crystal ......................................................................................... 24

2.1.4 The lattice factor for a crystal having an infinite number of unit cells .................. 26

2.1.5 Deviation of the lattice factor due to a finite size crystal ...................................... 26

2.1.6 Scattering vector $H$ ............................................................................................. 27

2.1.7 The Bragg conditions ............................................................................................. 28

2.1.8 Physical and experimental resolution limit ............................................................. 29

2.1.9 Structural heterogeneity and dynamic disorder ...................................................... 29

2.1.10 The mosaicity ....................................................................................................... 32

2.1.11 Lorentz factor ....................................................................................................... 33

2.1.12 Polarization factor ............................................................................................... 33

2.1.13 Total integral intensity scattered by a crystal ....................................................... 34
2.6 Experimental details to investigate the kinetic dose limit in PYP .......... 84
  2.6.1 Data collection ............................................................................. 84
  2.6.2 Absorbed dose calculation ......................................................... 85
  2.6.3 SVD analysis .............................................................................. 88

2.7 Experimental details to investigate the pH dependence of PYP photocycle
  89
  2.7.1 Data Collection ........................................................................... 89
  2.7.2 Kinetic analysis ........................................................................... 90
  2.7.3 Extrapolated maps ...................................................................... 90

2.8 Experimental details to investigate the temperature dependence of the PYP
  photocycle .......................................................................................... 91
  2.8.1 Time-resolved crystallography ......................................................... 91
  2.8.2 Time-resolved spectroscopy ......................................................... 93

3 RESULTS ......................................................................................... 97

3.1 Interaction of Nitric Oxide (NO) with Catalase: X-ray crystallographic
  analysis .................................................................................................. 97
  3.1.1 Crystal structures .......................................................................... 97
  3.1.2 Ammonia-bound Catalase form (Cat-NH$_3$) .................................. 99
  3.1.3 The five-coordinate Cat-5 .............................................................. 100
  3.1.4 Cat-NO Structure ........................................................................ 101

3.2 Interaction of Nitric Oxide (NO) with Catalase: Micro-spectrophotometric
  analysis .................................................................................................. 103
  3.2.1 Spectral changes for Cat-5, Cat-NH$_3$ and Cat-NO ......................... 103

3.3 Time-resolved Laue crystallography: Effect of X-ray dose on PYP ....... 105
  3.3.1 Data processing ............................................................................ 105
  3.3.2 Average absorbed dose ................................................................. 105
  3.3.3 Correction to the absorbed dose ($D_{corr}$) .................................... 106
  3.3.4 Correction for laser damage using the control experiment ............... 107
  3.3.5 First Right singular vectors (rSVs) from SVD ................................... 108
  3.3.6 The ratio of the amplitude to the offset .......................................... 109
  3.3.7 Relaxation times .......................................................................... 110

3.4 Time-resolved Laue crystallography: Effect of pH on PYP ............... 112
  3.4.1 Data processing ............................................................................ 112
  3.4.2 Averaged difference maps and extrapolated electron density maps of
    intermediates ..................................................................................... 114
  3.4.3 Structural analysis of intermediates at pH 7, pH 4 and pH 9 .............. 116
  3.4.4 Kinetic Analysis ........................................................................... 117

3.5 Effect of temperature on PYP photocycle ......................................... 117
  3.5.1 Time-resolved Laue crystallography ............................................. 117
3.5.2 Time-resolved spectroscopy ........................................................................ 120

4 DISCUSSION ........................................................................................................ 125

4.1 Catalase: structure and function ........................................................................ 125
  4.1.1 Significance for other heme proteins involved in electron transfer .......... 129
  4.1.2 Summary ........................................................................................................ 129

4.2 Time-resolved study on PYP ........................................................................... 130
  4.2.1 Effect of radiation dose ................................................................................ 130
  4.2.2 Effect of pH .................................................................................................... 134
  4.2.3 Effect of temperature .................................................................................... 138

4.3 Future applications .......................................................................................... 141

5 APPENDICES ......................................................................................................... 149

  5.1 Appendix A: Program 1 for time-resolved spectroscopic data collection 150
  5.2 Appendix B: Graphical user interface for the timing synchronization .... 153
  5.3 Appendix C: Program 2 for time-resolved spectroscopic data analysis ... 154
  5.4 Appendix D: Program 3 for the fitting of right singular vectors using a sum
                   of exponentials ............................................................................................ 157
  5.5 Appendix E: A figure showing the relaxation times obtained from time-
                   resolved crystallography ............................................................................. 159
# LIST OF FIGURES

**Figure 1.1** Chemical structure of the prosthetic group heme type ‘B’ ........................................................................................................................................6

**Figure 1.2** Quaternary structure of bovine liver catalase (BLC) ........................................................................................................................................7

**Figure 1.3 A.** Cartoon representation of the photoactive yellow protein .................................................................................................................................10

**Figure 1.4 A.** Photoactive yellow protein (PYP) photocycle ........................................................................................................................................12

**Figure 2.1 A.** Atomic form factors of oxygen, carbon and hydrogen atoms plotted as a function of sinθ/λ ..................................................................................24

**Figure 2.2 A.** An example of four unit cells with the translational vectors a, b, and c (not shown, lies parallel to axb) in real space ........................................................................25

**Figure 2.3** Representation of the diffraction geometry in reciprocal space for the conventional rotation method ........................................................................................................28

**Figure 2.4 A.** Several factors contribute to the static disorder and one of them is demonstrated .............................................................................................................................30

**Figure 2.5 A.** The atomic form factor of a carbon atom, normalized to its atomic number plotted against sin θ/λ and resolution d ................................................................................................................31

**Figure 2.6** Effect of the mosaicity χ on the reciprocal lattice points .....................................................................................................................................................32

**Figure 2.7** Schemes for different methods to crystallize proteins .......................................................................................................................................................35

**Figure 2.8 A.** A typical diffraction pattern obtained for catalase using monochromatic x-ray crystallography ..................................................................................................................36

**Figure 2.9** Scheme for monochromatic X-ray data collection .......................................................................................................................................................37

**Figure 2.10** Scheme for monochromatic X-ray data analysis .........................................................................................................................................................38

**Figure 2.11** A reciprocal space representation of Laue diffraction by a stationary crystal .................................................................................................................................43

**Figure 2.12** A typical λ-curve determined from an experimental data set collected at beamline BioCARS 14-ID-B at Advanced Photon Source (APS) ..........................................................................................45

**Figure 2.13** A scheme of time-resolved Laue crystallography .........................................................................................................................................................47

**Figure 2.14 A.** A pencil-shaped long PYP crystal mounted in a glass capillary for the time-resolved experiments ..........................................................................................................................49

**Figure 2.15 A.** The main steps associated with time-resolved Laue data processing ..................................................................................................................................................53

**Figure 2.16** Difference maps near the p-cinnamic acid (pCA) chromophore in PYP ........................................................................................................................................56

**Figure 2.17** A scheme of Singular value decomposition (SVD) .........................................................................................................................................................57

**Figure 2.18** Two examples of kinetic mechanisms used for slower part of the PYP photocycle .................................................................................................................................59

**Figure 2.19** Chromophore of photoactive yellow protein (PYP) .........................................................................................................................................................60

**Figure 2.20** Two types of samples used for the time-resolved spectroscopic experiments ..................................................................................................................................................61
Figure 3.12 A. Time-resolved difference absorption spectra from 30 μs – 2 s for PYP in solution at 295 K
........................................................................................................................................121

Figure 3.13 Time-resolved difference absorption spectra on crystalline slurry at different temperatures 122

Figure 4.1 Structural differences of the PYP chromophore in the pG and blue shifted intermediate state pB
........................................................................................................................................135

Figure 4.2 A schematic setup for a mix and inject time-resolved femtosecond (fs) serial crystallographic experiment at the XFEL.................................................................144
LIST OF TABLES

**Table 1.1** Noble prizes in the field of X-ray crystallography (Not a complete list) ........................................3

**Table 2.1** Data collection statistics of PYP at different temperatures. Table from Schmidt et al., 2013. ....93

**Table 2.2** Data collection parameters assigned to different time delays for time-resolved spectroscopic experiments ...................................................94

**Table 3.1** Data collection statistics for the three crystal forms of the catalase. Table from Purwar et al., 2011 ........................................ ...................................................97

**Table 3.2** Number of cavities in some selected proteins. Table from Purwar et al., 2011. ..................101

**Table 3.3** Tilt+bend angles for NO found in different heme-iron proteins with the iron in different oxidation states (Mb: myoglobin, NP4: nitrophorin 4). Table from Purwar et al., 2011. ........................102

**Table 3.4** Structural differences in Å relative to Cat-5 averaged over all 4 subunits. Table from Purwar et al., 2011 ...................................................103

**Table 3.5** Data statistics for 1st, 6th and 12th time-series after analysis with Precongnition/Epinorm. Table from Schmidt et al., 2012 ...................................................106

**Table 3.6** Absorbed doses (in 10⁵ Gy) for 12 consecutive time-series, before and after corrections .......108

**Table 3.7** D1/2 and D1/2K derived from uncorrected and corrected dose data. Table from Schmidt et al., 2012 ..........................................................111

**Table 3.8** Data collection statistics for PYP at three different pH conditions. Δρ denote highest positive and lowest negative difference electron density features in the unit of σ value that is used for the difference maps. Table from Tripathi et al., 2012 ...................................................112

**Table 3.9** Relaxation times from SVD analysis (τ1& τ2). Rate coefficients (k1, k2 & k3) and their inverse, τp from the posterior analysis. Table from Tripathi et al., 2012 ...................................................117

**Table 3.10** Relaxation times obtained (τi) from the fit of rSVs for all temperature settings .........118

**Table 3.11** Relaxation times (τi) and corresponding macroscopic rate coefficients (Λi) for PYP in solution and for a crushed single crystal. ..........................................................123

**Table 4.1** On and off rates as well as binding constants (K) for several protein-NO (or CO) complexes. Table from Purwar et al., 2011 ...................................................126

**Table 4.2** Calculated value of dose limits and employed dose rates for different type of proteins .........132

**Table 4.3** Rate of recovery from the signaling state pB to the dark state pG for PYP photocycle determined using time-resolved spectroscopy. The observed rates are compared with those reported by others .........139
ACKNOWLEDGEMENTS

First of all, I would like to thank my PhD supervisor Prof. Marius Schmidt for his great support, not only in scientific development and implementation of novel ideas but also for teaching me to lead a research project independently. Besides that, I would like to acknowledge him for reading and providing comments on my thesis. I would like to thank my committee members Prof. Abbas Ourmzad, Prof. Andy Pacheco, Prof. Dilano Saldin and Prof. Vali Raicu for giving important and useful comments on my thesis.

I thank my collaborators, Prof. Andy Pacheco for his valuable contribution in the catalase project and Dr. Vucika Srajer for her immense support during the experiments at synchrotron beamline. I would like to acknowledge them not only for their extraordinary help in the completion of the research projects but also for their friendly and supportive personality.

I would also like to thank my colleagues for a competitive discussion as well as creating a friendly atmosphere in the laboratories. I would like to acknowledge Jason Tenboer for his contribution in assembling the micro-spectrophotometer, Shailesh Tripathi for providing PYP crystals, Jennifer M. McGarry for her contribution in growing catalase crystals, Sahil Sangani and Lavanya Ramachandran. I thank all my friends for their constant support and encouragements.
Last but not the least, I am grateful to my husband, Vikas Malik for being supportive and encouraging throughout my PhD. I thank my family members for their care, love, motivation and support in difficult times. My deepest regard is to my parents and my brother for being a great source of inspiration and for their immense love and care.
1 Introduction

Proteins and nucleic acids are the most important biomolecules synthesized by living organisms. They are involved in different types of biological functions (Garrett & Grisham, 2009). For example, Deoxyribonucleic acids (DNA) and ribonucleic acids (Adams & Tsien, 1993) store and transmit essential hereditary information. Proteins are the building blocks of life and perpetuate almost all the necessary functions in a living organism. Protein folds into a unique structure which determines its function. Different types of proteins play various and distinct roles (Berg et al., 2001). For example, proteins like elastin, collagen and keratin provide structural support. Contractile proteins such as actin and myosin are involved in motion. Hemoglobin, found in red blood cells, is responsible for oxygen transport from the lungs to the body tissues. Antibodies are produced by the immune system when harmful substances such as bacteria and viruses are detected. Enzymes or bio-catalysts are the most important proteins. They are capable of catalyzing almost all the chemical reactions necessary for living beings (Cornish-Bowden, 2012).

1.1 Enzymes: structure and function

Enzymes are very specific and highly efficient. They catalyze particular reactions in a well-defined manner (Koshland, 1958). Malfunctions of enzymes may cause various types of disorders or diseases (Griffiths A.J.F, 1999, Copeland, 2000). If suitable drugs can be designed and developed, these diseases can be cured. Drug designing is possible only if the enzymes responsible for that particular disease are well understood. Enzymes are complex molecules and their specificity originate from their precise interaction with
their substrate (Koshland, 1958). This precision is a result of the unique three-dimensional structure of an enzyme. To get a deeper insight into the structural-functional relationship, it is crucial to determine the atomic structure of an enzyme. For three-dimensional structure determinations, X-ray crystallography is the most widely used technique.

1.2 Short history of crystallography

In 1912, Max Von Laue discovered that X-rays can be diffracted by crystals (Von Laue et al., 1913). With Bragg’s contribution, father and son, crystallography became a feasible technique for structure determination (Bragg, 1912). The first structure solved was that of simple inorganic salt (Bragg, 1913). In 1923, structures of a few important organic compounds such as hexamethylenetetramine were determined (Dickinson & Raymond, 1923). Several efforts were made to crystallize enzymes. Urease (EC 3.5.1.5) was the first enzyme to be successfully crystallized (Sumner, 1926). Nevertheless, it took more than 40 years from the first inorganic crystal structure to the first protein structure, sperm whale myoglobin (Kendrew et al., 1958). In 1963, the atomic structure of hemoglobin, which is four times larger than myoglobin, was solved (Muirhead & Perutz, 1963). Structures of these proteins provide an insight as to how respiratory systems work. In 1965, the first structure enzyme, lysozyme was obtained (Blake et al., 1965). Crystallography reached another level in 1984 when the structure of the first membrane protein, photosynthetic reaction center from Rhodopseudomonas viridis, was determined (Deisenhofer et al., 1984). The structure of the most important rotary molecular machine named as F1ATP-ase from bovine heart mitochondria was solved (Abrahams et al.,
Ribosome is one of the largest asymmetric complexes whose structure was solved in 2000 (Wimberly et al., 2000, Ban et al., 2000, Schluenzen et al., 2000). Soon after, the first structure of mammalian G-protein-coupled receptor (GPCR), from bovine rhodopsin was determined (Paczewski et al., 2000).

**Table 1.1 Nobel prizes in the field of X-ray crystallography (Not a complete list).**

<table>
<thead>
<tr>
<th>Year</th>
<th>Field</th>
<th>Name</th>
<th>Achievements</th>
</tr>
</thead>
<tbody>
<tr>
<td>1914</td>
<td>Physics</td>
<td>Max Von Laue</td>
<td>Diffraction of X-rays by crystals</td>
</tr>
<tr>
<td>1915</td>
<td>Physics</td>
<td>W. H. Bragg &amp; W. L. Bragg</td>
<td>Usage of X-rays for structure determination</td>
</tr>
<tr>
<td>1946</td>
<td>Chemistry</td>
<td>J. B. Sumner</td>
<td>Crystallization of first enzyme</td>
</tr>
<tr>
<td>1962</td>
<td>Chemistry</td>
<td>J. C. Kendrew &amp; M. Perutz</td>
<td>Three dimensional structures of globular proteins</td>
</tr>
<tr>
<td>1988</td>
<td>Chemistry</td>
<td>J. Deisenhofer, R. Huber &amp; H. Michel</td>
<td>Atomic structure of the first membrane protein</td>
</tr>
<tr>
<td>2009</td>
<td>Chemistry</td>
<td>V. Ramakrishnan, T. A. Steitz &amp; A. E. Yonath</td>
<td>High resolution three dimensional structure of ribosome</td>
</tr>
<tr>
<td>2012</td>
<td>Chemistry</td>
<td>R. J. Lefkowitz &amp; B. K. Kobilka</td>
<td>Structure of first G-protein-coupled receptor (GPCR)</td>
</tr>
</tbody>
</table>

In the 1980’s, time-resolved methods were introduced to the field of crystallography (Moffat, 1989). The technique is based on the Laue method which was originally discovered by Laue back in 1912, however employed by Moffat and coworkers 70 years later. Structural changes were investigated for different type of protein complexes such as Ha-Ras p21 during GTP hydrolysis (Schlichting et al., 1990) and myoglobin-CO complexes during photolysis (Srajer et al., 1996). Over the years, the technique kept developing and reached a matured state in the beginning of the 21st century. Recently, very early events of the trans to cis isomerization in a photoreceptor were successfully elucidated by employing this technique (Jung et al., 2013).
Apart from crystallography, other techniques such as nuclear magnetic resonance (NMR) spectroscopy (Wuthrich, 1990) and electron microscopy (Unwin & Henderson, 1975) are also used to obtain atomic structures. All the atomic structures determined from any of these techniques are stored in the Protein Data Bank (PDB) (Berman et al., 2000). Currently, about 90,000 PDB entries are available, of which 80,000 are solved using crystallographic techniques. Less than 10,000 structures have been solved using either NMR or electron microscopy. This shows that X-ray crystallography has been a major technique for the determination of the three dimensional structures of bio-molecules.

1.3 From static structure to function

Static atomic structures of the macromolecules are obtained from conventional X-ray crystallography. The static information can provide a significant insight about the function; however, for a deeper understanding, one needs to watch a protein in action. For functional investigations of proteins, time-resolved techniques are exquisitely suitable. When an enzyme catalyzes a reaction, several intermediates may form and decay along the reaction pathway. Molecular interactions and motions that occur on a fast time scale can be relatively easily visualized with time-resolved spectroscopic techniques such as fluorescence (Holzwarth, 1995, Millar, 1996), infrared (Schotte et al., 2003, Brudler et al., 2001) and absorption spectroscopy (Topp et al., 1971, van Amerongen & van Grondelle, 1995, van Stokkum et al., 2004). From spectroscopic techniques, one can determine a plausible kinetic mechanism of a catalytic reaction. These spectra may also provide some limited structural information. A technique that combines kinetics and structure determination is time-resolved crystallography (Moffat, 1989).
During the time course of a reaction, subsequent structural changes can be probed with time-resolved crystallographic methods. In a reaction, some of the intermediate states are occupied for a very short period of time and are therefore difficult to detect at ambient temperature. These short-lived states are called metastable states or transient states. It is possible to probe these states, if the life time of the intermediates can be extended for a longer period of time either by lowering the temperature or by adjusting other parameters such as pH (Hajdu et al., 2000, Moffat & Henderson, 1995, Stoddard, 2001). For example, the structures of short-lived intermediates involved in the ligand binding to myoglobin are determined using these trapping approaches (Schlichting et al., 2000, Teng et al., 1994). The disadvantage of trapping techniques is that they may perturb the kinetic mechanism (Moffat & Henderson, 1995). The goal is to study a reaction that evolves freely at physiological temperatures without any perturbation. With modern synchrotron X-ray sources and fast detectors, it is possible to probe intermediates on the order of nanoseconds and faster even at room temperature.

1.4 Catalase

Heme proteins are the most abundant proteins found in living organisms (Bertini I., 1994). They are involved in a number of different biological functions (Berg et al., Paoli et al., 2002). As mentioned above, hemoglobin and myoglobin are responsible for oxygen transport (Antonini & Brunori, 1971, Jain & Chan, 2003). Proteins like cytochrome c oxidase (Wikstrom, 1977) behave as an electron carrier whereas catalase protects the living cells from oxidative damage originating from excess of reactive oxygen species (Deisseroth & Dounce, 1970). These proteins contain heme (Fig. 1.1) in their active sites,
which is required for their biological function (Berg et al., 2006). This prosthetic group consists of a heterocycle ring known as protoporphyrin IX with an iron (Fe$^{2+}$) atom at the center (Fig. 1.1). When a ligand is bound to the heme protein, even slight structural changes in the heme group can significantly affect the protein activity (Paoli et al., 2002). These proteins are excellent for functional study especially if crystallographic methods are used. With the ligand binding approach, not only the function of heme proteins can be understood but also the results can be generalized to different types of proteins.

As mentioned above, catalase (E.C. 1.11.1.6) is a metalloenzyme with heme in its active site. It is found in many bacteria and almost all plants and animals (Murthy et al., 1981). Catalase catalyzes the disproportionation of toxic hydrogen peroxide ($\text{H}_2\text{O}_2$) into oxygen and water (Eq. 1.1)

\[
\text{H}_2\text{O}_2 \rightarrow \frac{1}{2} \text{O}_2 + \text{H}_2\text{O}
\]
This reaction is very important to all aerobically respiring organisms. Reactive oxygen species like $O_2^{2-}$ and $O_2^-$ (superoxide) are inevitable byproducts of aerobic respiration, necessitating specialized enzymes for their elimination (Aebi, 1984, Halliwell & Gutteridge, 1990, Michiels et al., 1994). Typical examples for mammalian catalase are bovine liver catalase, BLC (Ko et al., 1999) and human erythrocyte catalase, HEC (Ko et al., 2000).
BLC was crystallized for the first time in 1937 (Sumner & Dounce, 1937) and its crystal structure was determined in 1981 (Murthy et al., 1981). The quality of atomic models have been subsequently improved (Reid et al., 1981, Ko et al., 1999). BLC is a tetramer (Fig. 1.2A) (Tanford & Lovrien, 1962, Sund et al., 1967). Each of the four monomers (MW~60 KDa) consist of 507 amino acid residues and contains a heme group (Fig. 1.2B) with the iron in the ferric state (Fe$^{3+}$) (Torii et al., 1970, Stern, 1936, Schroeder et al., 1982).

Several mechanisms have been proposed for BLC catalysis (Fita & Rossmann, 1985, Alfonso-Prieto et al., 2007). Among these, the Fita-Rossman model is the most widely accepted. Based on this model, the reaction involves a transfer of a total of four electrons in two major steps. First, two electrons are transferred from the porphyrin complex to one molecule of H$_2$O$_2$ and then two electrons are accepted from a second H$_2$O$_2$ molecule (Eq. 1.2 & 1.3). Compound 1 forms in the first part of the reaction (Eq. 1.2) and attacks another H$_2$O$_2$ molecule abstracting a hydride ion (H$^-\$) and leaving a hydroxyl ion (OH$^-\$) at the iron position (Fita & Rossmann, 1985). Iron itself becomes Fe(III) and one electron is transferred back to the porphyrin. A proton temporarily stored at the N atom of His 74 (Fig. 1.2B) is then transferred to the OH$^-\$ at the iron site. This leads to the formation of a water molecule, which is very weakly bound to the iron and can be quickly released (Eq. 1.3).

$$\text{porphy} - \text{Fe (III)} + \text{H}_2\text{O}_2 \rightarrow \underbrace{\text{porphy}^{**}}_{\text{compound 1}} - \text{Fe (IV)} = \text{O} + \text{H}_2\text{O} \quad (1.2)$$

$$\text{porphy}^{**} - \text{Fe (IV)} = \text{O} + \text{H}_2\text{O}_2 \rightarrow \text{porphy} - \text{Fe (III)} + \text{H}_2\text{O} + \text{O}_2 \quad (1.3)$$
Within the active enzyme, the Cat-H$_2$O$_2$ adduct, the so called compound 1 (see above), is not sufficiently long-lived to be structurally characterized. Various types of ligands such as azide (Murshudov et al., 2002), formate (AndreOLETTI et al., 2003), oxo (compound I and II) (Gouet et al., 1996, Alfonso-Prieto et al., 2007), cyanide and 3-amino-1,2,4-Triazol (Putnam et al., 2000) can be bound to the heme iron. These active site’s adducts are prepared to gain an insight into the structural information about the intermediates crucial to the catalytic mechanism. If nitric oxide (NO) can be successfully bound to the catalase, the Cat-NO adduct may mimic compound 1.

NO as a ligand has been bound to the other heme proteins such as leghemoglobin Lb (Rohlfs et al., 1988), Myoglobin Mb (Rohlfs et al., 1988, Laverman et al., 2001) and nitrophorin NP (Andersen et al., 2000). Several complexes have been developed and implemented for a controlled release of NO in biological systems (Maragos et al., 1991). Despite this fact, the structure of the Cat-NO complex has never been characterized before because nitric oxide tends to dissociate rapidly from the crystals. Here, we aimed at the structure determination of the nitrosylated-BLC complex using monochromatic crystallography.

1.5 PYP

As described above, heme proteins such as catalase are ideal for static structure determinations. Ligand binding to these proteins can provide an insight into the function of proteins. Our goal, however, is to investigate the dynamic behavior of the biomolecules in general. We need to develop techniques that can be applied to various types of biological systems. For this, a well-understood model system is required. Photoactive
yellow protein (PYP) provides such a model reaction that evolves with several intermediates at ambient temperatures. Apart from that, PYP crystals of small size and high stability make this protein an excellent model system for investigating functional behavior. PYP and its photocycle are now described in detail.

![Figure 1.3 A. Cartoon representation of the photoactive yellow protein (in violet). The para-coumaric acid chromophore (pCA; colored green) is embedded into the protein. B. The chromophore and its nearby residues that play a crucial role in the photocycle. Panel B from Jung et al., 2013.](image)

PYP (Meyer, 1985) is a water-soluble protein with a molecular weight of 14kDa. It is initially isolated from the bacterium *Halorhodospira* (formerly *Ectothiorhodosopira*) *halophila*. The X-ray structure of PYP (Fig. 1.3A) (Borgstahl et al., 1995) shows that it shares a common structural motif with the typical PAS domain family (Meyer, 1985, Pellequer et al., 1998). The para-coumaric acid (pCA) chromophore (Fig. 1.3) is embedded into this domain and is believed to play a key role in the negative phototactic effects (Sprenger et al., 1993) in the various photosynthetic bacteria (Kort et al., 1996). The chromophore is covalently attached to Cys69 through a thioester bond and forms
hydrogen bonds with amino acid residues Tyr42, Glu46 and with backbone amide of Cys69 (Fig. 1.3B) (Van Beeumen et al., 1993).

Upon absorption of a blue photon, PYP enters into a fully reversible photocycle with several intermediates ranging from picoseconds to seconds (Fig. 1.4). The photocycle has been studied extensively with various time-resolved techniques such as UV-Vis absorption spectroscopy (Meyer et al., 1989, Hoff et al., 1994, Imamoto et al., 2001, Ujj et al., 1998), Fourier transform infrared spectroscopy (Brudler et al., 2001, Imamoto et al., 2001, van Wilderen et al., 2006), NMR spectroscopy (Rubinstenn et al., 1998), resonance Raman spectroscopy (Pan et al., 2004) and Laue crystallography (Genick et al., 1998, Schmidt et al., 2004, Ren et al., 2001, Ihee et al., 2005, Rajagopal et al., 2005, Schotte et al., 2012, Jung et al., 2013). In some cases, it is possible that PYP may not successfully enter into the photocycle even after absorbing a blue light photon (Groot et al., 2003, van Wilderen et al., 2006). As a result, a ground state intermediate (GSI, Grey part of Fig. 1.4A) forms on the femtosecond time scale and decays to the ground state (pG) in ~6 ps. The GSI was revealed by ultrafast infrared spectroscopy; however, the structure of this intermediate is still elusive because time-resolved crystallographic methods at ambient temperatures are lacking the time resolution that is required to capture the GSI intermediate.

The nomenclature used for the intermediates associated with this photocycle (Fig. 1.4A) is the same as used in earlier studies (Jung et al., 2013, Genick et al., 1997a, Ihee et al., 2005, Kim et al., 2012). When a protein crystal is illuminated with a photon of 485 nm, an electron jumps to an excited state. A part of the energy is immediately dissipated
Figure 1.4 A. Photoactive yellow protein (PYP) photocycle. Solid black curves: part of the photocycle used to investigate the effect of pH on the structures, Red dashed curve: part of the photocycle used to examine the impact of radiation damage on the protein crystal. Figure from Schmidt et al., 2012 B. A photocycle consistent with WT-PYP time-resolved electron-density maps. Carbon atoms of the refined pG, I₁, Iₐ₁, pR₁, pR₂ and pB intermediate structures are shown in grey, orange, cyan, dark blue, green, and purple, respectively. Oxygen, sulfur and nitrogen atoms are shown in red, yellow and blue, respectively. The arrows indicate large atomic movements from one intermediate to the next. Figure from Jung et al., 2013.
(Martin et al., 1983) and the remaining energy triggers the photo-isomerization of the chromophore. A theoretical calculation performed on WT-PYP and its mutant, E46Q, revealed that the earliest intermediate $I_T$ forms in less than 100 ps (Groenhof et al., 2004). The structure of $I_T$ (Fig. 1.4B) was recently determined by time-resolved crystallography (Jung et al., 2013, Schotte et al., 2012). It resembles the spectroscopically observed intermediate $I_o$ (Fig. 1.4A) (Imamoto et al., 2001, Unno et al., 2004). This intermediate is not fully transformed into the cis conformation and exists in a highly strained state (Schotte et al., 2012, Jung et al., 2013, van Stokkum et al., 2004). The remaining isomerization of the $I_T$ cannot occur through the standard mechanism (Liu & Asato, 1985), because the chromophore is buried within the host protein and therefore interacts with surrounding side chains (Borgstahl et al., 1995). This severely restricts the conformational mobility of the chromophore and imposes spatial constraints on it (Liu & Asato, 1985, Warshel & Barboy, 1982). Therefore, the isomerization pathway proceeds through a volume-conserving mechanism (Müller et al., 1998, Liu et al., 2007). The distorted intermediate $I_T$ bifurcates into two structurally distinct cis intermediates $I_{CT}$ and pR1 (Fig. 1.4B). $I_{CT}$ forms in 1.7 ns via a bicycle-pedal (BP) mechanism (Xie et al., 1996, Genick et al., 1998, Ren et al., 2001, Groot et al., 2003, Ihee et al., 2005, van Wilderen et al., 2006, Imamoto et al., 2002) whereas pR1 forms via a hula twist (HT) mechanism (Liu & Asato, 1985, Andresen et al., 2005) with a life time of ~3ns. The mixture of these two cis-intermediates is similar to the spectroscopically observed intermediate $I_o^#$ (Fig. 1.4A). In $I_{CT}$ the carbonyl-oxygen is flipped to the other side but the chromophore head is fixed by hydrogen bonds to Tyr42 and Glu46. In pR$_1$ the chromophore head hydroxyl loses one hydrogen bond. The highly strained intermediate $I_{CT}$ relaxes to pR$_2$. Both
intermediates pR1 and pR2 form in parallel (Tripathi et al., 2012). The pR1 and pR2 resemble the intermediates pR_{E46Q} and pR_{CW} observed from the nanosecond crystallographic studies, respectively (Ihee et al., 2005, Rajagopal et al., 2005). In pR1, the hydrogen bonds to the Tyr 42 and backbone amide of Cys69 are maintained and that to the Glu46 is broken (Fig. 1.4B). On the other hand, the chromophore in pR2 maintains the hydrogen bonds to Tyr42 and Glu46 and loses the hydrogen bond to the backbone amide of Cys69. The mixture of pR1 and pR2 is equivalent to spectroscopic intermediate, pR (Fig. 1.4A), which absorbs in the range of 465 nm whereas the dark state absorbs at 446nm (Ujj et al., 1998, Hoff et al., 1994). Both intermediates pR1 and pR2 ultimately decay in ~200 µs to the long lived signaling state pB, which absorbs at 355nm (Fig. 1.4) (Schmidt et al., 2004, Ihee et al., 2005). The chromophore head forms new hydrogen bonds with the displaced Arg52 and with an additional water that appears near the entrance to the chromophore pocket (Fig. 3.8 I.G). The phenolate oxygen and Arg52 are exposed to the solvent. At the end of the photocycle, this blue shifted intermediate reverts to its initial dark state on the millisecond time scale. When the protein structure relaxes back to the dark state, all the hydrogen bonds are re-established. The system has been thoroughly studied; however, it is unknown how this protein reacts at different parameters such as pH, temperature and X-ray dose. We investigate the effects of these parameters on the protein kinetics using time-resolved crystallographic methods.

### 1.6 Effect of X-rays doses on PYP

In X-ray crystallography, a protein crystal is repeatedly exposed to X-ray radiation during data collection. These exposures can potentially damage the protein and affect its
dynamics (Dubnovitsky et al., 2005, Adam et al., 2004, Purwar et al., 2011, Schlichting et al., 2000). At cryogenic temperatures, the dose limit up to which meaningful data can be collected is also known as Henderson’s limit (Henderson, 1990). Below a threshold of approximately $2 \times 10^7$ Gy, the damages are not substantial because up to this limit the damage varies linearly with X-ray dose (Teng & Moffat, 2000). Above this limit, damages start to accumulate and can cause excessive damage to the protein crystals. In order to reduce the radiation damage, different types of techniques such as cryo-cooling (Kuzay et al., 2001, Nicholson et al., 2001) and the usage of a free-radicals scavenger (Murray & Garman, 2002) have been developed. At room temperature, protein crystals are even more susceptible to radiation damage. The damaging effects of X-rays on the kinetics of a protein have never been quantified before. Apart from X-rays, laser pulses employed in a time-resolved experiment (Moffat, 1989) may also destroy the protein crystals. Here, we performed time-resolved crystallographic experiments to assess the impact of not only X-ray doses but also laser pulses on the structural and kinetic analysis of PYP. This investigation shows how many data sets can be collected from a single crystal of PYP without affecting the kinetics.

### 1.7 Effect of pH on the PYP

The pCA chromophore is de-protonated and present in its anionic form when PYP is in the ground state (Baca et al., 1994, Borgstahl et al., 1995, Kim et al., 1995). Upon entering into a photocycle (Fig. 1.4A), the net uptake of one proton is involved during the life time of the intermediate pB (Meyer et al., 1993). The protonation of the chromophore may be responsible for the blue-shift of ~95 nm of the pB state (~355 nm) with respect to
the pG state (~446 nm) (Fig. 4.1). The transfer of protons takes place either from the Glu-46 (Baca et al., 1994, Hendriks et al., 1999) or from the solvent (Borucki et al., 2002, Genick et al., 1997b). Protonation and deprotonation of the chromophore as well as nearby amino acid residues may lead to pH- or salt-dependent protein activity (Genick et al., 1997b, Hoff et al., 1997, Harigai et al., 2003). For native PYP, the photo bleaching which corresponds to the pR → pB transition is slightly accelerated at low pH conditions (Genick et al., 1997b, Demchuk et al., 2000). On the other hand, recovery to the dark state (pB → pG) is the fastest at pH 8 and decreases for all other pH conditions (Genick et al., 1997b, Demchuk et al., 2000). These studies show that the pH has a pronounced effect on the kinetics of the photocycle. Although it has been widely investigated using spectroscopic techniques, the effect of pH on the structures of reaction intermediates on the atomic length scale is still elusive. In this thesis, time-resolved crystallography is employed to investigate these effects.

1.8 Effect of temperature on PYP
In addition to the external parameters such as X-ray dose and pH, temperature also affects the kinetics of PYP. Several time-resolved spectroscopic studies have been performed on PYP at one (Meyer et al., 1987, Genick et al., 1997b, Imamoto et al., 1996, Ujj et al., 1998, Unno et al., 2002, Meyer et al., 1989) or at a few different temperature settings (Hellingwerf et al., 2003, Van Brederode et al., 1996). For example, picoseconds time-resolved crystallographic measurements were recently performed on PYP at room temperature (Jung et al., 2013, Schotte et al., 2012). This ps study thoroughly describes the isomerization pathways and provides structural information of earlier intermediates.
A detailed investigation how the temperature affects the structure and kinetics of the PYP photocycle has never been performed. Here, we employ time-resolved crystallography to investigate the photocycle rigorously at 14 different temperature settings starting from -40 °C to 70 °C (Schmidt et al., 2013).

The functional study at different temperatures can also be performed using time-resolved spectroscopy. This enables us to investigate the protein kinetics with two different but widely used techniques. The results obtained from the two complementary techniques can be compared to explore the similarities and differences in the kinetic behavior. These experiments have never been previously accomplished due to unavailability of a suitable micro-spectrophotometer that can probe the spectroscopic changes on fast time scales in a temperature-controlled environment. In our lab, we designed a fast spectrophotometer fulfilling all of the aforementioned requirements. This spectrophotometer is employed to investigate the PYP photocycle at different temperatures. The function of a protein may be different in solution as compared to its crystalline phase (Yeremenko et al., 2006). To determine these differences, we also performed time-resolved spectroscopic experiments on PYP in solution.

The rates of chemical reactions depend on the temperature. Their temperature dependence can be described using Van't Hoff-Arrhenius equation, which is given by

\[ \Lambda = \nu \exp\left(-\frac{E_a}{k_B T}\right) \]  

(1.4)

where \( \Lambda \) is the macroscopic rate coefficient, \( \nu \) is the prefactor, \( E_a \) is the energy of activation, \( k_B \) is the Boltzmann factor and \( T \) is the temperature. Eyring (Eyring, 1935) tied this equation to a transition state at top of the barrier of activation. This is
represented as a transition state equation (Eq. 4.1), which is further described in the discussion. With this information, extracting the energetics of a biomolecular reaction has been a topic of interest and has also been challenging. So far only limited information is available (Takeshita et al., 2002). Here, we show how thermodynamic properties such as enthalpy, entropy and Gibbs free energy of the barrier can be extracted solely from the temperature dependent crystallographic data.

References


2 Material and methods

Monochromatic X-ray crystallography is the major technique to determine the three dimensional structures of bio-molecules. To further understand the function of these molecules, time-resolved Laue crystallography is employed. As a comparison, kinetic analyses are performed with time-resolved absorption spectroscopy on solution as well as with crystals. These techniques are described in detail in the following sections.

2.1 Monochromatic X-ray crystallography

2.1.1 Scattering from an atom

To investigate the molecule on the atomic length scale, the wavelength of the employed radiation should be comparable to the size of the atoms or to the length of a chemical bond. Therefore, X-ray radiation with wavelengths in the range of 1 ~ 2 Å is an excellent choice. When X-rays are incident on an atom, they are scattered by electrons. The scattering power of an atom is given by the atomic form factor \((f)\). \(f\) is proportional to the number of electrons in an atom. Therefore, scattering from a heavy atom such as iron is much stronger than the scattering from a carbon, oxygen or a nitrogen atom. The atomic form factor also depends on the scattering angle \((2\theta)\) and decreases significantly at higher values of \(2\theta\) (Fig. 2.1A).

2.1.2 Scattering from a molecule

All atoms in a molecule scatter according to their atomic form factors. The superposition of all the scattered waves results in the structure factor of a molecule which is given by

\[
F_M^M(H) = \sum_{j=1}^{N} f_j e^{2\pi i H \cdot \vec{R}_j}
\]  

(2.1)
where \( \mathbf{q} \) is the scattering vector, \( N \) is the total number of atoms in a molecule, \( f_j \) is the atomic form factor of the \( j^{th} \) atom, \( \mathbf{r}_j \) is the position vector of the \( j^{th} \) atom. An example for a molecule having 4 atoms with their atomic form factors and corresponding phases is drawn in an Argand diagram (Fig. 2.1B). The resultant structure factor of this molecule is

\[ F_M = \sum_{j=1}^{N} f_j e^{i \phi_j} \]

\[ \phi_j = \phi_{1j} + \phi_{2j} + \phi_{3j} + \phi_{4j} \]

**2.1.3 Scattering from a crystal**

Due to the weak interaction of X-rays with matter, the intensity scattered from a molecule is very weak. An ensemble of molecules is required to intensify scattering. In a single crystal, molecules are arranged in a periodic lattice. The structure factor of a crystal is the superposition of the scattered waves from all atoms in the crystal which can be written as

\[ F_{\text{crystal}} = \sum_{\mathbf{g}} e^{i \mathbf{g} \cdot \mathbf{R}} \]

where \( \mathbf{g} \) are the reciprocal lattice vectors and \( \mathbf{R} \) are the direct lattice vectors.
where $M$ is the total number of atoms in the crystal, $f_m$ is the atomic form factor of the $m^{th}$ atom and $\mathbf{r}_m$ is the position vector of the $m^{th}$ atom (Fig. 2.2A). If a crystal has $N$ unit cells with each of them containing $J$ atoms, the total structure factor of the crystal can be separated in two terms which can be written as

$$S = \sum f_m \exp \left( i \mathbf{g} \cdot \mathbf{r}_m \right) + \sum f_j \exp \left( i \mathbf{g} \cdot \mathbf{r}_j \right)$$

where $\mathbf{g}$ is the lattice vector and $\mathbf{r}_j$ is the position vector of $j^{th}$ atom in the unit cell (Fig. 2.2A). The first term is known as the lattice factor $G$ and the second term represents the structure factor of a molecule.

**Figure 2.2 A.** An example of four unit cells with the translational vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ (not shown, lies parallel to x) in real space. $\mathbf{r}_m$ (in red) denotes the position vector of the $m^{th}$ atom with respect to a common origin $O$. $\mathbf{r}_j$ can be represented as a sum of lattice vector $\mathbf{m}^* \mathbf{a}^*$ (in blue) and the position vector $\mathbf{R}_j$ (in green) with respect to the origin $O'$. **B.** The representation of a real space and the associated reciprocal space. $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ denote real space vectors (in black) and $\mathbf{a}^*$, $\mathbf{b}^*$, and $\mathbf{c}^*$ denote corresponding reciprocal lattice vectors (in red). Three parallel lattice planes of miller index (110) are shown in blue (dashed line). Scattering vector $\mathbf{H}$ is represented as linear combination of reciprocal space vectors with integer multiple $h = 1$, $k = 1$ and $l = 0$. 
2.1.4 The lattice factor for a crystal having an infinite number of unit cells

If a crystal has U, V and W unit cells in the direction of translation vectors $\vec{a}$, $\vec{b}$, and $\vec{c}$, respectively then the lattice factor $G$ (Eq. 2.3) is expressed as a triple sum

$$G = \sum_{u=0}^{U} e^{2\pi i u \vec{H} \cdot \vec{a}} \cdot \sum_{v=0}^{V} e^{2\pi i v \vec{H} \cdot \vec{b}} \cdot \sum_{w=0}^{W} e^{2\pi i w \vec{H} \cdot \vec{c}}$$

where $u$, $v$ and $w$ are the integer numbers. For an infinitely large crystal, this triple sum is zero unless the scalar products of the scattering vector and the lattice vectors are integer numbers.

$$\vec{H} \cdot \vec{a} = h$$

$$\vec{H} \cdot \vec{b} = k$$

$$\vec{H} \cdot \vec{c} = l$$

Here $h$, $k$ and $l$ are integer numbers and these relations are known as the Laue conditions. This shows that a crystal scatters X-rays only in those directions for which the Laue conditions are satisfied. The lattice factor $G$ is then given by

$$G = U \cdot V \cdot W = N$$

where $N$ is the total number of unit cells. The structure factor amplitude of the crystal is amplified by $N$ due to the lattice factor contribution.

2.1.5 Deviation of the lattice factor due to a finite size crystal

Crystal dimensions are not infinite. Their finite sizes cause certain deviations from the Laue conditions. The intensity amplification due to the lattice factor $G$ is given by

$$G^2 = \frac{\sin^2 U \pi \vec{H} \cdot \vec{a}}{\sin^2 \pi \vec{H} \cdot \vec{a}} \cdot \frac{\sin^2 V \pi \vec{H} \cdot \vec{b}}{\sin^2 \pi \vec{H} \cdot \vec{b}} \cdot \frac{\sin^2 W \pi \vec{H} \cdot \vec{c}}{\sin^2 \pi \vec{H} \cdot \vec{c}}$$
The total area calculated under each of these functions shows that the scattered intensity is amplified by \( N \) rather than \( N^2 \).

### 2.1.6 Scattering vector \( \vec{H} \)

A crystal consists of lattice lines and lattice planes. When X-rays hit these lattice planes, they are scattered such that the incident and the scattered beam make an equal angle to the planes. The direction of the scattering vector \( \vec{H} \) is perpendicular to these reflecting planes. To represent the scattering vector \( \vec{H} \), a coordinate system in reciprocal space is constructed, in which the Laue conditions are automatically fulfilled (Fig. 2.2B). The reciprocal lattice vectors are calculated using the real space basis vectors as follows

\[
\vec{a}^* = \frac{\vec{b} \times \vec{c}}{\vec{a} \cdot \vec{b} \times \vec{c}}; \quad \vec{b}^* = \frac{\vec{c} \times \vec{a}}{\vec{a} \cdot \vec{b} \times \vec{c}}; \quad \vec{c}^* = \frac{\vec{a} \times \vec{b}}{\vec{a} \cdot \vec{b} \times \vec{c}}; \tag{2.8}
\]

where \( \vec{a}^* \), \( \vec{b}^* \) and \( \vec{c}^* \) are reciprocal lattice vectors and \( \vec{a} \), \( \vec{b} \) and \( \vec{c} \) are real space vectors. The scattering vector \( \vec{H} \) can be represented as a linear combination of the reciprocal lattice vectors which is given by

\[
\vec{H} = h\vec{a}^* + k\vec{b}^* + l\vec{c}^* \tag{2.9}
\]

where \( h \), \( k \) and \( l \) are the integer numbers. For a lattice plane of Miller index \((h, k, l)\), the length of scattering vector \( \vec{H} \) is given by the following equation

\[
|\vec{H}| = \sqrt{(h\vec{a}^*)^2 + (k\vec{b}^*)^2 + (l\vec{c}^*)^2} = \frac{1}{d_{hkl}} \tag{2.10}
\]

where \( d_{hkl} \) is the interplanar distance of a plane \((h, k, l)\). All the higher order planes \((nh, nk, nl)\) share a common miller index \((h, k, l)\). For these set of planes, the direction of scattering vector \( \vec{H} \) is the same, however the length is different which is given by
where \( n \) is the integer multiple and \( \mathbf{s} \) is defined above. As a consequence, the scattering vector is a crystal property.

**Figure 2.3** Representation of the diffraction geometry in reciprocal space for the conventional rotation method. An Ewald sphere of radius \( 1/\lambda (\text{Å}^{-1}) \) is drawn where \( \lambda \) is the wavelength of monochromatic X-rays. The origin of the reciprocal lattice is at point \( O \). Black dots, arranged in a periodic manner, denote the reciprocal lattice points (RLPs). The directions of incoming and scattered beams are represented by \( \mathbf{s}_1 \) and \( \mathbf{s}_0 \), respectively. The lengths of vector \( \mathbf{s}_1 \) and \( \mathbf{s}_0 \) are the same due to the elastic scattering condition. \( 2\theta \) represents the scattering angle and \( \mathbf{s} \) is the scattering vector. RLP in blue intersects with the Ewald sphere and satisfies the diffraction condition. A sphere of radius \( 1/d_{\text{min}} \) denotes the experimental resolution limit. RLPs (in green) do not diffract because they lie outside this sphere of radii \( 1/d_{\text{min}} \).

### 2.1.7 The Bragg conditions

Bragg conditions arise from the combination of Laue conditions and elastic scattering conditions. Due to the elastic scattering of X-rays, the wavelength of incoming ray...
that of scattered ray $\vec{s}_1$ are same as shown in the Fig. 2.3. The relationship between scattering angle and scattering vector is given by

$$\sin \theta = \frac{|\vec{H}|/2}{1/\lambda}$$  \hspace{1cm} (2.12)

Using Eq. (2.11) and (2.12), the well-known Bragg condition is derived:

$$2d_{hkl} \sin \theta = n\lambda$$  \hspace{1cm} (2.13)

### 2.1.8 Physical and experimental resolution limit

All the reciprocal lattice points (RLPs) for which the Bragg condition is satisfied contribute to a diffraction pattern. However, the RLPs for which the scattering vectors $\vec{H}$ are larger than the diameter of the Ewald sphere ($2/\lambda$) never penetrate the Ewald sphere. These lattice points lie outside the physical resolution limit $d$ which is given by

$$d = \frac{\lambda}{2}$$  \hspace{1cm} (2.14)

where $1/\lambda$ is the radius of the sphere centered at the origin $O$. Protein crystals do not diffract to physical resolution limit due to certain experimental limitations. This enforces an experimental resolution limit which is given by $d_{\text{min}}$ (Fig. 2.3). Consequently, only RLPs lying inside the resolution sphere of $1/d_{\text{min}}$ can contribute to the diffracted intensity.

### 2.1.9 Structural heterogeneity and dynamic disorder

There are two types of disorders found in the protein crystals: static and dynamic. In static disorder, no motion is observed on the time scale of the experiment. In an ideal crystal, unit cells are repeated in a regular manner in three dimensions. All the molecules in each unit cell have the same positions as well as the same orientations. However, in a real crystal, molecules or parts of the molecules do not occupy the same position and also
may not have exactly the same orientation. This is one of the factors which contribute to the static disorder. In addition, proteins are large molecules and their structures are quite flexible. This causes structural heterogeneity in which atoms are slightly shifted in each unit cell. The conformation of a molecule can be different from one unit cell to the other (Fig. 2.4A). The disorders described above are of static type. In the case of dynamic disorder, motions are much faster than the time scale of the experiment. Atoms in the crystal vibrate about their equilibrium position and therefore, at any instance of time do not occupy identical position in the different unit cells. As a consequence, an average over all the atomic positions is obtained which is equivalent to an enlarged electron density at a fixed atomic position (Fig. 2.4B). Additionally, molecules might fluctuate between different conformations. These factors give rise to dynamic disorder in a crystal.

**Figure 2.4 A.** Several factors contribute to the static disorder and one of them is demonstrated. Four different orientations of a molecule in four unit cells are shown. The resultant conformation (red ellipsoid) is extended spatially. **B.** A factor causing dynamic disorder is shown. A molecule is vibrating about an equilibrium position (red spheres). An average over seven atomic positions (partially overlapped spheres) is equivalent to an enlarged electron density of the molecule (red ellipsoid). Modified version of the figure from http://www.christian.naether.uni-kiel.de/pdf.
Each of these different types of disorder, static and dynamic, contribute to the Debye Waller Factor (DWF)

\[
\text{DWF} = \frac{\langle x^2 \rangle}{\lambda^2} \sin^2 \theta
\]

where \( \langle x^2 \rangle \) is the mean square displacement, \( \theta \) is half of the scattering angle, \( \lambda \) is the wavelength of the monochromatic radiation and \( \langle x^2 \rangle \) is the B-factor. The value of DWF is less than one except in the forward direction (\( \theta = 0 \), Eq. 2.15). This diminishes the scattering intensity for any scattering angle larger than zero. Higher angles are significantly more affected as compared to lower angles (Fig. 2.5A). Since the DWF is parameterized by the mean square distribution of the atomic positions, it does distinguish between static and dynamic disorder. The types of disorders may be differentiated if

**Figure 2.5 A.** The atomic form factor of a carbon atom, normalized to its atomic number plotted against and resolution \( d \). The effect of an increase in the mean displacement on the atomic form factor is also shown. **B.** An oversimplified conceptual schematic showing the effect of lowering the temperature on static and dynamic disorders. At ambient temperature, overlapping red and blue ellipsoids indicate that static disorder (red ellipsoid) and dynamic disorder (blue ellipsoid) cannot be distinguished. When temperature is decreased, static disorders are unaffected whereas dynamic disorders start to decrease. Modified version of the figure from http://www.christian.naether.uni-kiel.de/pdf.
diffraction data are collected at different temperatures. At lower temperatures dynamic motion seems to reduce significantly whereas static disorders should be unaffected (Fig. 2.5B). X-ray crystallographic experiments are performed at cryogenic temperatures. This substantially diminishes the effects of dynamic disorders on the diffraction data.

### 2.1.10 The mosaicity

Perfect crystals generate sharp reciprocal lattice points whose intersection with the Ewald sphere give rise to intense Bragg peaks. Real crystals are not perfect and contain various defects. Due to these imperfections, it can be considered that a crystal is composed of small mosaic blocks, whose orientations are slightly different from each other. As a consequence, a given Bragg reflection is observed through an angular range which is known as the mosaicity ($\chi$). $\chi$ determines the width of the so called rocking curve (Fig. 2.6). The rocking curve also broadens with the increase in the divergence and the bandwidth of the employed X-ray radiation. If a crystal is rotated with a rotation step $\Delta\phi$ smaller than the mosaicity, the reflections are only partially recorded in one diffraction

**Figure 2.6** Effect of the mosaicity $\chi$ on the reciprocal lattice points. An Ewald sphere of radius $1/\lambda$ is drawn. $1/N$ represents the natural line width (in blue) of a reflection where $N$ is the total number of unit cells in the crystal. The broadening of the reciprocal lattice points due to mosaicity are shown in red double lines.
pattern. The remaining parts of the reflections are recorded in subsequent images. For a highly mosaic crystal, reflections start to overlap. This may affect the quality of the obtained intensities.

2.1.11 Lorentz factor

When a crystal rotates, the associated reciprocal lattice and their lattice points rotate through an Ewald sphere. A reflection remains in the diffracting position for a certain period of time given by

$$\frac{\lambda}{\omega \sin 2\theta}$$

(2.16)

where $\lambda$ is wavelength of the incoming X-ray beam, $\omega$ is the angular velocity of rotation and $2\theta$ is the scattering angle. In addition, the scattered intensity depends on the intersection of the reciprocal diffraction volume with the Ewald sphere. The ratio of this intersected area and the total surface area of an Ewald sphere is dependent on $1/\lambda^2$. After combining both aforementioned effects, the Lorentz factor contributes a $\lambda^3$ dependence to the scattered intensity.

2.1.12 Polarization factor

When an electron interacts with electromagnetic radiation such as X-rays, the electron accelerates and radiate like an oscillating dipole. Since no scattering is observed in the direction of the electron propagation, the scattered intensity must be corrected. This correction factor is known as the polarization factor. For the unpolarized beam, the scattered intensity is reduced by the factor $\frac{1+\cos^2 2\theta}{2}$. On the contrary, synchrotron beam is strongly polarized in the horizontal direction. If this polarized beam is used to perform an experiment, a crystal is rotated such that its rotation axis is aligned along the polarization
direction of the beam. This maximizes the scattering intensity perpendicular to the axis of rotation.

### 2.1.13 Total integral intensity scattered by a crystal

After incorporating all the factors described previously, the scattering intensity from a crystal is given by

$$I_{\text{int}}(hkl) = \frac{\lambda^3}{\omega} \cdot N \cdot \sigma_e \cdot I_0 \cdot \frac{1 + \cos^2 2\theta}{2\sin 2\theta} \cdot |F^M(hkl)|^2$$  \hspace{1cm} (2.17)

where $I_{\text{int}}$ is the integral reflection intensity, $\lambda$ is wavelength of incoming X-rays, $N$ is the total number of unit cells, $\sigma_e$ is the Thomson scattering cross section of an electron, $I_0$ is the intensity of the incident X-rays, $2\theta$ is scattering angle, $|F^M|$ is the structure factor amplitude of a molecule at the position $h$, $k$ and $l$. Due to the $\lambda^3$ dependence (Eq. 2.17), the radiation with a longer wavelength such as X-ray from Cu-K$_{\alpha}$ (1.54 Å) is an excellent choice for in-house protein crystallography, where $I_0$ is comparatively small.

### 2.1.14 Sample preparation and mounting

A crystal is picked up by a cryo-loop from the hanging drop (Fig. 2.7A) or sitting drop (Fig. 2.7B) of a crystallization tray. The crystal is then soaked in a cryoprotectant, a buffer that typically contains sugar, polyethylene glycol and glycerol, for 15 ~ 25 s to reduce ice formation upon its freezing. In some cases, crystals are very sensitive. Such crystals are first soaked in a stabilization buffer for ~20 s and then in the cryoprotectant to reduce the osmotic shock. Crystals treated with cryoprotectant are mounted in a loop and shock frozen at 100K using a cryostream. If the goal is to bind a small molecule to the crystal, the crystal is soaked in a solution containing that molecule before applying the cryoprotectant. After that, it is looped and frozen as described above.
Figure 2.7 Schemes for different methods to crystallize proteins. Circle in white shows a single well of the 24-well plate. PR: protein solution, R: reservoir solution, W: water diffusing from protein solution to reservoir solution P: precipitant diffusing from reservoir to protein solution. High vacuum grease (green ellipses) is used to seal the cover slide. A. A wall plate (in grey) is used for the hanging drop method. The protein droplet (colored yellow) is placed upside down on the glass cover slide to allow equilibration between the protein and the reservoir solution. B. A wall plate used for the sitting drop method is shown in grey. The protein solution (yellow colored) is placed within a dip in the upward direction.

2.1.15 Experimental setup

All the monochromatic crystallographic experiments described in this thesis are performed at beamline BioCARS 14-BM-C at the advanced photon source (APS). A typical experimental set up of the beamline is described in the following. X-rays are generated by inserting bending magnets in the path of a high energy electron beam. A Germanium (111) single crystal is used as a monochromator (http://biocars.uchicago.edu/page/14-bm-c-beamline). X-rays of wavelength 0.9 Å (E = 13.8 keV) with ΔE/E = 3.1 x 10^-4 are used. The beam is focused to a typical spot size of 100 x 300 μm^2 at the sample position. The flux of X-ray beam in this spot size is ~5 x 10^{11} photons/s. This flux is used to illuminate the crystal during the data acquisition. A charge coupled device (CCD) area detector ADSC Q315 is used to record X-ray photons scattered from a crystal.
**2.1.16 Data collection**

The crystal is initially exposed to the X-rays in a random orientation. A typical diffraction pattern obtained for catalase to resolution 1.9Å is shown in Fig. 2.8A. The crystal is then rotated through a rotation range $\phi$ (Fig. 2.9A) in steps of $\Delta\phi$. The diffraction patterns obtained after each rotation step are collected. $\phi$ usually depends on the symmetry of a crystal. A smaller $\phi$ can be selected for a crystal of higher symmetry while a crystal having no symmetry must be rotated through 180° (Fig. 2.9B). $\Delta\phi$ is chosen based on the mosaicity of the crystal and its typical value is 0.5° ~ 1°.

**2.1.17 Data analysis and refinement**

All the steps associated with the data processing are shown in Fig. 2.10. Programs such as Mosflm (Leslie, 2006) and HKL2000 (Otwinowski & Minor, 1997) are widely used to reduce monochromatic data. However, we used Mosflm, because it provides the best...
Figure 2.9  Scheme for monochromatic X-ray data collection. A. A crystal (yellow colored) is illuminated by X-rays and rotated about the $\varphi$-axis (in red) during data collection. The area detector CCD is shown in grey. B. The diffraction patterns are shown from $\varphi = 0^\circ$ to $180^\circ$ with $\Delta \varphi = 0.5^\circ$.

statistics for protein crystals with large unit cells and with overlapping reflections. The quality of data statistics obtained from Mosflm is shown in Tab. 3.1. Three main steps are involved in data reduction with Mosflm: Indexing, Refinement and Integration. In indexing, crystal parameters such as unit cell dimensions and the initial crystal orientations are determined. In the next step, these two crystal parameters as well as the mosaicity are refined. Detector parameters (detector position and detector orientation) and beam parameters (beam orientation and beam divergence) are also refined. For each reflection, two masks are defined such that one mask covers the Bragg peak and another one covers the background. Background-subtracted reflection intensities are determined by integrating the photon counts found in the individual pixels. Intensities are less accurate for weak and overlapping reflections. Profile fitting is used to determine the intensity of such reflections. Sample profiles are first derived from well separated and non-overlapping reflections and then fitted to the reflections with lower intensities.
The volume calculated under two dimensional fitted profile yields the integrated intensity on an arbitrary scale. A common scale factor is required to bring all the reflection intensities on the same level. For that, intensities are assumed on the same scale for an individual diffraction pattern as well as for few subsequent ones. The intensities of the partials from the subsequent images are summed up with the CCP4 program SCALA.
(CCP4, 1994, Evans, 2006). The summed intensities from the partials and those from the full reflections are used to determine the scale factor for each individual diffraction pattern. By applying these scale factors, the reflection intensities from all the diffraction patterns are brought to the same scale. Symmetrically equivalent reflections as well as reflections with multiple measurements are merged and averaged to obtain unique reflections. From a list of unique reflections intensities, structure factor amplitudes and their standard deviations are calculated using the CCP4 program TRUNCATE (CCP4, 1994)(Fig. 2.10). The quality of reduced data is determined by examining the $R_{\text{merge}}$, $I/\sigma_I$ and completeness. The $R_{\text{merge}}$ denotes the agreement between multiple observations of a given reflection.

$$R_{\text{merge}} = \frac{\sum_{hkl} \sum_{j=1}^{N} |I_{hkl}(j) - \bar{I}_{hkl}|}{\sum_{hkl} \sum_{j=1}^{N} I_{hkl}(j)} \quad (2.18)$$

where $N$ is a number of multiple observations for the reflection $(h, k, l)$ and $\bar{I}_{hkl}$ is the average intensity of this reflection. For large unit cells, overall $R_{\text{merge}}$ values smaller than 0.08 ~ 0.1 are acceptable. $I/\sigma_I$ denotes the ratio of intensity to their standard deviation which is equivalent to the signal to noise ratio. This value should be larger than two in the highest resolution shell. Completeness is a measurement of the coverage of the reciprocal space in a full data set and should be more than 75 ~ 80 %.

### 2.1.18 Molecular replacement and crystallographic refinement

Once structure factor amplitudes are determined, phases are required to calculate the electron density $\rho$. At coordinates $X, Y$ and $Z$ in the unit cell, $\rho$ can be represented as

$$\rho(XYZ) = \frac{1}{V} \sum_{h} \sum_{k} \sum_{l} |F(hkl)| e^{i\alpha_{hkl}} \cdot e^{-2\pi i(hX+kY+lZ)} \quad (2.19)$$
where $V$ is the volume of the unit cell, $|F(hk l)|$ and $\alpha_{hkl}$ are structure factor amplitude and phase of the reflection (h, k, l), respectively. Phases can be determined using three major approaches: multiple isomorphous replacement (MIR), multiple anomalous dispersion (MID) and molecular replacement (MR). If a protein structure that has a sequence similar to the protein under investigation is available, then molecular replacement can be used. Otherwise, one has to follow the complicated process of ab-initio phase determination with MIR or MAD. The molecular replacement method involves two important steps. The first step is rotation in which the orientation of a known atomic model is found in the crystal. The next one is translation in which the positions of correctly orientated search model are found in the unit cell. If the known structure has been determined for the same crystal form as the protein to be investigated, then rotation and translation steps are not required anymore. The known atomic model can directly be used as a starting model. However, refinements are necessary to increase the agreement between the calculated and the observed structure factor amplitudes (Fig. 2.10). Before any of the refinement steps are performed, all the water molecules and unwanted cofactors are removed from the initial atomic model. In our case, the program Crystallographic & NMR systems (CNS) (Bruenger et al., 1988) is used for the model refinement. As a first step, the orientation of a protein molecule within the unit cell is refined by a rigid body refinement at low resolution such as 3 Å (Fig. 2.10). All the subunits of a protein molecule are treated as rigid domains. To correct the large structural errors in the atomic model simulated annealing refinement also known as molecular dynamics refinement is performed. The temperature of a molecule is first raised sufficiently high and then the molecule is allowed to cool down slowly. Typically, a
2000K protocol to full resolution is applied for large protein molecules. After that, conventional positional refinement is performed to optimize the atomic positions. The residual between calculated \( F_{hkl}^{\text{calc}} \) and observed \( F_{hkl}^{\text{obs}} \) structure factor amplitudes is minimized in a least square method (Eq. 2.20).

\[
\sum_{hkl} w\left( |F_{hkl}^{\text{obs}}| - |F_{hkl}^{\text{calc}}(X,Y,Z,B)| \right)^2 + \sum_{i=1}^{\text{restraints}} (r_i^{\text{obs}} - r_i^{\text{calc}})^2 \rightarrow \text{minimum} \quad (2.20)
\]

The first term in this minimization protocol is called X-ray term or X-ray residual \( Q_{\text{xray}} \), in which atomic coordinates X, Y, Z and B-factors are free parameters. \( w \) is the weighting factor for this term. In order to increase the ratio of observables to free parameters, stereochemical parameters such as bond lengths, bond angles, dihedral angles, van der Waals contacts and planar groups are used. If small deviations are allowed for these parameters during the refinement, they are known as restraints. The sum of least square residuals of these restraints from its target values is known as geometrical term or geometrical residual (Eq. 2.20). Both X-ray and geometric terms are minimized simultaneously until convergence is reached. B-factors of all the atoms are determined using restrained B-factor refinement with a 20 ~ 40 step protocol. After the refinement steps, SigmaA-weighted \( 2mF_{\text{obs}} \) \( DF_{\text{calc}} \) maps are generated (Rupp, 2010). These Fourier maps are inspected using a molecular viewer program Xfit from XtalView package (McRee, 1999).

### 2.1.19 Water search and addition

For a complete atomic structure, crystallographically visible water molecules are found and added to the atomic model of the protein. Water molecules up to a certain sigma (\( \sigma \)
level are searched from a difference electron density map ($F_{\text{obs}} - F_{\text{calc}}$). These water molecules are added to the structure. However, if the binding of a small molecule to a protein is under investigation, one should avoid adding water molecules to the protein active site. After insertion of all the water molecules, atomic positions are refined using conventional refinement. Once generated, electron density map $2F_{\text{obs}} - F_{\text{calc}}$ and difference map $F_{\text{obs}} - F_{\text{calc}}$ are inspected. In the $F_{\text{obs}} - F_{\text{calc}}$ map, positive electron density may indicate the positions of additional water molecules or that of ligands expected to be bound. The negative electron density may denote badly refined regions. If the electron density map reveals misorientations of amino acid residues, they are manually reoriented into the electron density map using Xfit (McRee, 1999). Positional refinement is performed again after these model corrections. Both of these steps are repeated until no obvious misorientations are detected.

### 2.1.20 Structure validation

Once the atomic model is properly refined, the residual between observed and calculated structure factor amplitudes can be determined. This residual is known as R-factor, which can be written as

$$R = \frac{\sum_{hkl} \{ |F_{\text{obs}}(hkl)| - |F_{\text{calc}}(hkl)| \}}{\sum_{hkl} |F_{\text{obs}}(hkl)|}$$

where $|F_{\text{obs}}(hkl)|$ and $|F_{\text{calc}}(hkl)|$ are the observed and calculated structure factor amplitudes, respectively. An atomic model whose R-factor is less than about 0.25 can be considered reliable. A Ramachandran plot can also be used to judge the geometric quality of an atomic model. This plot shows the possible peptide bond angles ($\phi$, $\psi$) of the protein main-chain. Unrealistic values can be identified for an incorrect atomic model. In
addition, Luzzati plots (Drenth, 1999) are used to estimate the coordinate error of the calculated model by plotting the R-factor as a function of resolution.

2.2 Laue and time-resolved X-ray crystallography

2.2.1 Polychromatic X-rays

In the Laue technique, a single crystal is illuminated with polychromatic X-rays to obtain a diffraction pattern (Fig. 2.8B). Due to the broad bandwidth of these X-rays, a wide

Figure 2.11 A reciprocal space representation of Laue diffraction by a stationary crystal. A reciprocal lattice with the origin at O is shown. Two Ewald spheres of radii $1/\lambda_{\text{min}}$ and $1/\lambda_{\text{max}}$ are drawn where $\lambda_{\text{min}}$ and $\lambda_{\text{max}}$ indicate the minimum and the maximum wavelength of the polychromatic X-rays, respectively. Basis vectors of reciprocal lattice are represented as $\mathbf{a}$ and $\mathbf{b}$. Blue colored reciprocal lattice points (RLPs) lie between both the Ewald spheres and therefore satisfy the diffraction condition. Corresponding reflections are obtained on the detector with scattering angle $2\theta$. RLPs shown in black are not in the reflecting positions. Red colored RLPs (-1, 2, 0) and (-2, 4, 0) generate harmonically overlapping reflections. RLPs (in green) do not diffract because they lie outside a sphere of radii $1/d_{\text{min}}$ where $d_{\text{min}}$ represents the maximum resolution limit.
range of the reciprocal lattice is covered. The minimum ($\lambda_{\text{min}}$) and the maximum ($\lambda_{\text{max}}$) wavelength limit of the polychromatic radiation gives rise to two Ewald spheres of radii $1/\lambda_{\text{min}}$ and $1/\lambda_{\text{max}}$ (Fig. 2.11). The diffraction condition is satisfied for all those reciprocal lattice points (RLPs) which lie between both the Ewald spheres.

### 2.2.2 Instantaneous recording of integral intensity

The use of polychromatic radiation provides an opportunity to collect integrated reflection intensities instantaneously without rotating the crystal. As described in monochromatic technique (Sec. 2.1.11), the $\lambda$-dependence of the Lorentz factor comes from the rotation. This term is absent in Laue method. As a consequence, Laue scattering intensity depends only on $\lambda^2$.

### 2.2.3 Spatially overlapping reflections

A Laue diffraction pattern is typically very crowded with reflections as compared to that obtained from the monochromatic method. Consequently, a large number of spatially overlapping reflections is present. This overlap occurs due to a very small angular separation of the adjacent diffracted X-ray beam. The density of overlapping spots depends on the scattering angle $2\theta$ as well as on the geometrical conditions such as the unit cell parameters (Cruickshank et al., 1991). The spatial overlap can be resolved by using analytical and numerical reflections profiles. These profiles are derived from the well-separated, non-overlapping reflections and then used to separate these spatial overlaps (Ren & Moffat, 1995).

### 2.2.4 $\lambda$-curve for wavelength normalization
The intensity of the polychromatic radiation varies as a function of wavelength. Therefore, reflections are excited by different frequencies and intensities depending on the orientation of the crystal (Schmidt, 2008, Ren et al., 1999). These reflection intensities must be brought to a common scale. This procedure is known as wavelength normalization that results in a so called \( \lambda \)-curve (Srajer et al., 2000). The \( \lambda \)-curve can be determined directly from the data set of original and unscaled reflection intensities using single order Bragg reflections also known as singlets. This curve corrects for all wavelength-dependent parameters such as the source spectrum, detector response, the overall absorption correction and scattering from a crystal. If the X-ray spectrum of employed X-rays is not stable, the determination of the \( \lambda \)-curve is difficult. A typical \( \lambda \)-curve of full width half maximum (FWHM) of 0.057 Å obtained from a synchrotron X-ray source as shown in Fig. 2.12.

### 2.2.5 Harmonically overlapping reflections
When Bragg’s law is applied, many orders of each Bragg reflections such as \((d, \lambda), (d/2, \lambda/2), \ldots (d/n, \lambda/n)\) may be stimulated simultaneously. All of them are scattered at the same angle and therefore exactly superimpose on the diffraction pattern (Fig. 2.11). This is called harmonic or energy overlap. For example, reflections \((2\ 4\ 6)\) and \((4\ 8\ 12)\) generate harmonics and share a common index \((1\ 2\ 3)\). The harmonic overlap mainly applies to reflections at low resolution. If these reflections are dismissed from the analysis, the completeness of the low resolution data is reduced. This effect is referred as the ‘low resolution hole’. To resolve the harmonic overlap, a particular harmonic and its symmetry mates are measured multiple times at different crystal orientations (Ren & Moffat., 1995). The \(\lambda\)-curve must be already determined as described in the Sec. 2.2.4.

As an example, here we show how to resolve a harmonic in which three orders of reflections are superimposed. The observed intensity of a given harmonic \(I_1^{obs}\) at different crystal orientations can be written as linear combination of individual intensities

\[
\begin{align*}
I_1^{obs} &= \lambda_{11}I_1 + \lambda_{12}I_2 + \lambda_{13}I_3 \\
I_2^{obs} &= \lambda_{21}I_1 + \lambda_{22}I_2 + \lambda_{23}I_3 \\
I_3^{obs} &= \lambda_{31}I_1 + \lambda_{32}I_2 + \lambda_{33}I_3
\end{align*}
\]  

(2.22)

where \(I_{i}^{obs}\) is the intensity of a given harmonic at \(i^{th}\) orientation, \(I_i\) is the individual reflection intensity of \(i^{th}\) order reflection and \(\lambda_{ii}\) are the component of \(\lambda\)-curve, which is known. Eq. 2.22 can be rearranged in a matrix representation

\[
I^{obs} = \Lambda I
\]

(2.23)

where \(I^{obs}\) are observed intensities of a particular harmonic, \(\Lambda\) can be determined from the \(\lambda\)-curve. Individual intensities \(I\) can be calculated by matrix inversion.

### 2.2.6 Scattering background
A large fraction of the polychromatic X-rays does not diffract into Bragg reflections. This fraction generates scattering background. This background could be due to various factors such as crystal disorder, scattering by the capillary, scattering by air and Compton scattering. As a consequence, it is difficult to determine very weak reflection intensities that usually appear at high resolution. The background can be significantly reduced by using radiation with smaller bandwidth. A highly intense X-ray beam of a narrow band pass (Graber et al., 2011) can be generated by using an undulator. One has to ensure that the beam is large enough to cover the band width accepted by each reflection. With this undulated radiation, data collection at higher resolution becomes possible and reflection intensities can be measured with higher accuracy (Srajer et al., 2000).

2.2.7 The mosaicity
The Laue method is very sensitive to the crystal mosaicity. If the mosaicity is increased, Laue diffraction spots become elongated. In most cases, the elongation of these spots further increases when a reaction evolves in a protein crystal. Visual inspection of the spots-elongation in the Laue patterns may be a potential method to check whether reaction initiation is successful or not. Higher mosaicity causes additional spatial overlap in an already crowded Laue image. Therefore, crystals with the mosaicity in the range of 0.2 ~ 0.4° should be used for time-resolved experiments (Schmidt *et al*., 2005a, Ihee *et al*., 2005).

### 2.2.8 Time-resolved Laue method

The use of the entire bandwidth of X-ray reduces the exposure time by several orders of magnitude as compared to the monochromatic method (Srajer *et al*., 2000). A Laue image can be collected on the order of 100 picoseconds. Such short exposure times make this technique an excellent choice to investigate the dynamics of bio-molecules (Amorós *et al*., 1975, Cruickshank *et al*., 1987). In time-resolved investigations, a reaction is first initiated and then subsequent changes are probed with polychromatic X-rays at different time delays ($\Delta t_i$) (Fig. 2.13). The time resolution of these experiments depends on the duration of the laser pulse or the X-ray pulse, whichever is longer. Reaction initiation can be achieved by various methods such as a rapid change in the concentration of substrates, cofactors, protons or electrons or by a rapid change in the temperature or pressure. However, the photo-activation of a chromophore molecule embedded within the protein is the most convenient one and widely used. For example, the photoreceptor PYP can be easily activated on the order of picoseconds and even higher. .
2.2.9 Sample mounting

Time-resolved diffraction experiments are performed at ambient temperature. A protein crystal is mounted in a glass or quartz capillary of 1mm diameter (Fig. 2.14A). A small drop of stabilization buffer is placed at the end of the capillary to keep the crystal moist (Schmidt et al., 2005a). If the crystal is dry, it degrades faster during its repetitive exposure to the consecutive laser pulses. On the other hand, a crystal embedded into too much liquid can easily move from its mounting position after exposure to a laser pulse. The capillary is mounted on the goniostat in a direction perpendicular to both the X-ray pulse and the laser pulse (Fig. 2.14A).

![Figure 2.14 A. A pencil-shaped long PYP crystal mounted in a glass capillary for the time-resolved experiments. Black solid line represents the total translation of the crystal during the entire data collection. The rotation axis of the crystal is shown in green. Laser spot size (blue) and X-ray spot size (red) are shown. Figure from Schmidt et al., 2010. B. Geometry of PYP crystal settings while illuminated by X-ray and laser pulses. The X-ray beam (red ellipsoids) probes the volume near the surface of the crystal that is illuminated by the laser light. The diameter of the laser beam (blue) is 200 µm, which is substantially larger than the horizontal component of the X-ray beam (h = 90 µm). This facilitates the alignment procedure. Arrow on the left denotes positive y’-displacements of the crystal. A part of elliptical region (in orange) shows the new position of X-ray beam after y’-displacement. The crystal translation along the long axis is 440 µm. The length, leff = 530 µm is used for the dose calculation. t is the thickness of the crystal. Dashed box: approximation of the beam with a rectangular box, dotted box: new position of the X-ray beam after displacement y’. Figure from Schmidt et al., 2012.](image-url)
2.2.10 Experimental setup

All time-resolved crystallographic experiments described in this thesis are performed at beamline BioCARS 14-ID-B at the Advanced Photon Source (APS). A typical experimental setup is described in the following. A pulse from an Nd:YAG pumped laser (Opolette HE II) is used to initiate the reaction in the crystal. The full width half maximum (FWHM) of the pulse is about 4 ns. It is transported through a tapered fiber and focused to ~200 µm spot at the sample position. The measured energy density of the laser pulse at the sample is 4 mJ/mm². After reaction initiation, subsequent structural changes that may range from nanosecond to second are probed by polychromatic X-rays. Short and intense X-ray pulses of 100 ps are generated from an undulator. The average wavelength ($\lambda_{\text{ave}}$) of the X-ray radiation is 1.05 Å (Fig. 2.11) which corresponds to an average energy ($E_{\text{avg}}$) of 11.8 keV. The bandwidth $\Delta E_{\text{avg}}/E_{\text{avg}}$ is 10%. Each X-ray pulse contains $3.5 \times 10^{10}$ photons (Graber et al., 2011). The X-ray pulse is focused at the sample position to a typical spot size of 90 µm (h) X 60 µm (v), where h and v denote the full width half maximum (FWHM) of the pulse in the horizontal and vertical direction, respectively (Fig. 2.14A). The Laser pulse and the X-ray pulse are orientated perpendicular to each other as well as to the capillary in which the crystal has been mounted (Sec. 2.2.9). In most cases, protein crystals are exquisitely optically thick and therefore, the laser pulse penetration into the crystal is shallow (Fig. 2.14B). This results in a reaction initiation primarily close to the illuminated crystal surface. The crystal is positioned in the X-ray beam such that only this surface layer is probed by X-rays (Schmidt et al., 2012). To determine the precise position of this surface layer, the crystal is translated in the vertical direction across the X-ray beam and several diffraction
patterns are collected. Once the X-ray beam crosses the edge of the crystal, diffraction pattern starts to fade. By analyzing these patterns, the edge can be determined. With this, the overlap of X-ray beam with the laser illuminated volume can be maximized (Fig. 2.14B). This procedure is known as edge scan (Schmidt et al., 2012, Graber et al., 2011).

2.2.11 Data collection

A time range that covers a reaction from the nanoseconds to the end is selected. Time delays ($N_t$) are equally distributed on the logarithmic scale with typical 3 time points per logarithmic decade. Once a time series is collected from one particular crystal orientation, the crystal is rotated by an angle $\Delta \theta$ for another orientation. These orientations are selected such that they are maximally spaced across reciprocal space. The advantage of this protocol is that if an experiment terminates prematurely, the collected data don’t have any preferred orientation. One can calculate a meaningful electron density map even from an incomplete data set. The angular range $\Delta \theta$ depends on the bandpass of the employed radiation (Ren et al., 1999). The values of $\Delta \theta$ between 2 ~ 3° are usually sufficient. Each time when a crystal is reoriented, it is also translated along its long axis so that a fresh volume can be exposed (Fig. 2.14B). In this type of data collection, time acts as a fast variable and orientation as a slow one. For a complete data set, Laue images at several different orientations ($N_o$) of the crystal must be collected. Usually, the number of photons in a single pulse even from the strongest X-ray source is not enough to produce a sufficiently strong diffraction pattern. Three to seven X-rays exposures ($N_p$) are required per diffraction pattern for an acceptable signal to noise ratio depending on the crystal size. A typical Laue diffraction pattern obtained for PYP with 4 X-ray exposures is
shown in Fig. 2.8B. Between two exposures, one should wait for enough time so that the sample can revert to its initial state. The waiting time depends on the velocity of the reaction. For a fast reaction, the waiting time can be short. A reference or dark data without laser illumination must be collected for each crystal to minimize the crystal specific errors.

### 2.2.12 Data reduction by Precognition

The goal of data reduction is to determine the accurate reflection intensities from the raw Laue diffraction images (Fig. 2.15A). For that, a software package *Precognition* and *Epinorm* (RenzResearch) is used. The first steps are indexing and integration, which are performed with *Precognition*. The final steps is scaling for which *Epinorm* is used.

Indexing is the assignment of the Miller indices to all spots on a Laue image. Indexing is typically required only for one image because angular relations between the indexed and all other patterns are known. An initial estimate of the unit cell parameters and crystal orientations is obtained. The best orientation can then be selected based on number of matched spot as well as on the root mean square deviations (RMSD). With this selected orientation, the displacements between the observed and predicted spots in the indexed pattern are then minimized by adjusting the geometrical parameters such as unit cell and detector parameters. This process is known as geometrical refinement. The RMSD and the number of matched spots are the two criteria, which are inspected to judge the quality of refined data. Once the indexed image is properly refined, its best orientation is distributed to all other images. These images are refined with respect to the first indexed image. The next step is to integrate the Laue spots in the images. Before we do that, overlapping spots must be separated.
Figure 2.15 A. The main steps associated with time-resolved Laue data processing. Steps shown in pink boxes are performed by available software (Schmidt et al., 2003 & Schmidt et al., 2004). B. Analysis of time-resolved spectroscopic data. A matlab script for data processing and the Matlab fitting routine are shown in the appendices.
For example, spatially overlapped reflections are deconvoluted using analytical profile fitting as discussed in the Sec. 2.2.3. On the other hand, resolving harmonically overlapped spots (Sec. 2.2.5) is not straightforward. As mentioned, a lambda curve (Sec. 2.2.4 & Fig. 2.12) is first generated and then applied to separate these spots. Integrated intensities for the resolved diffraction spots are then obtained.

2.2.13 Scaling by Epinorm

Integrated reflection intensities from diffraction images corresponding to the different orientations are on different scale. Epinorm is used to bring them on the same scale. For that, an initially estimated λ-curve is used for wavelength normalization (Sec. 2.2.4). Scaling parameters are then applied to these normalized integrated intensities. Once all the reflection intensities are brought to the same scale, the final step is to merge all the redundant and symmetry-related intensities. As a result, a list of unique reflection intensities is obtained, from which structure factor amplitudes can be determined.

2.2.14 From structure factor amplitudes to electron density maps

After data processing, structure factor amplitudes of the initial dark state $|F_{obs}^D(hkl)|$ and a corresponding set of time dependent structure factor amplitudes $|F_{obs}(hkl,t)|$ are obtained (Fig. 2.15A). To generate electron density maps, not only structure factor amplitudes but also phases are required. For time-resolved experiments, the dark state model must be available. From this model, the calculated structure factor amplitudes $|F_{calc}^D(hkl)|$ with their phases $\phi_{calc}^D$ can be obtained. Observed structure factor amplitudes of dark $|F_{obs}^D(hkl)|$ are scaled to the calculated one $|F_{calc}^D(hkl)|$ and hereby brought to the absolute scale (Schmidt et al., 2010). The time-dependent observed
amplitudes $|F_{obs}(hkl,t)|$ are then scaled to the reference data in order to bring them also on the absolute scale. Difference structure factor amplitudes $\Delta F_{obs}(hkl,t)$ for each time point are calculated by subtracting $|F_{Dobs}(hkl)|$ from the $|F_{obs}(hkl,t)|$. The difference amplitudes $\Delta F_{obs}(hkl,t)$ corresponding to each time point are multiplied by a weighing factor $w$, which is given by

$$w = \frac{1}{1 + \frac{\sigma^2}{\langle \sigma^2 \rangle} + \frac{\Delta F(hkl,t)^2}{\langle \Delta F(hkl,t)^2 \rangle}}$$

(2.24)

where $\sigma$ is the standard deviation of $|\Delta F_{obs}(hkl,t)|$. This ensures that the observations with large uncertainties and those with large differences in structure factor amplitudes are down-weighted for the difference map calculation. The weighted difference structure factor amplitudes are normalized by the average weighting factor to keep them on the absolute scale. Phases $\phi_{calc}^D$ calculated from the dark state model are combined with weighted difference structure factor amplitudes $w|\Delta F_{obs}(hkl,t)|$. Weighted time-dependent difference electron density $\Delta \rho(t)$ are calculated as

$$\Delta \rho(t) = \frac{1}{V_e} \sum_{hkl} w|\Delta F_{obs}(hkl,t)| e^{i\phi_{calc}^D(hkl)} e^{-2\pi i(hX+kY+lZ)}$$

(2.25)

where $X$, $Y$, and $Z$ are components of the position vector in the coordinate system of unit cell, $h$, $k$, and $l$ are indices of Laue reflections, $V_e$ is the volume of unit cells, $w$ is the weighting factor of the difference structure factor amplitudes $|\Delta F_{obs}(hkl,t)|$. Electron density on an absolute scale directly relates to the occupancy or fractional concentration if the structural differences are small and noise is moderate. This is referred to as the
difference approximation. The obtained electron density maps can be inspected with the program Xfit as also mentioned in the monochromatic method. A typical difference electron density map is shown in Fig. 2.16. Positive and negative electron density peaks are observed in these difference maps. Negative electron density features represent the loss of electrons and account for those atoms which were present at a given position in the dark model and have moved away from their positions to occupy new positions. On the other hand, the positive electron density features denotes the gain of electrons and account for the new positions occupied by atoms.

2.2.15 Singular value decomposition of difference electron density map

The time courses of weighted difference electron density maps are analyzed by singular value decomposition (SVD) using the program SVD4TX (Schmidt et al., 2003, Zhao & Schmidt, 2009).

Figure 2.16 Difference maps near the p-cinnamic acid (pCA) chromophore in PYP. The dark state model is shown in yellow, red and blue representing carbon, oxygen and nitrogen atoms, respectively. Positive electron density features: blue or cyan, negative electron density features: red or white. Contour levels: red/white (-3σ/-4σ), blue/cyan (+3σ/+4σ) A. Dark subtracted from intermediate state pR. B. Dark subtracted from intermediate pB.
In the difference electron density map, a confined region is specified for the SVD analysis because other parts of the map do not contain valuable information and add only noise to the analysis. First, those regions are selected for the mask where large changes in the electron density map are expected. The mask is then evolved by allowing only those grid points where the difference electron density is above or below a certain $\sigma$ level that occur at least in one time point (The $\sigma$ value represents the root mean square deviation of the difference electron density maps from the mean value determined for all the grid points). The masked data is arranged in a data matrix ($A$) which is then decomposed by SVD into left singular vectors (columns of $U$), right singular vectors (columns of $V$) and corresponding singular values (diagonal elements of $S$) such that $A = USV^T$ (Fig. 2.17).

The left singular vectors (ISVs) contain the time independent structural information. The temporal variations of these ISVs are contained in corresponding right singular vectors (rSVs). Singular values obtained from the diagonal matrix imply the number of significant rSVs. Typically, the first couple of rSVs are significant and all other contain noise. For example, if two significant singular vectors are present, the first ISV represents...

**Figure 2.17** A scheme of Singular value decomposition (SVD). The data arranged in a matrix $A_{mn}$. The matrix $A$ is decomposed into the matrix $U_{mn}$ containing left singular vectors, the diagonal matrix $S_{nn}$ having diagonal elements and the matrix $V_{nn}^T$ with right singular vectors.
an average of the difference electron density maps and first rSV describes its time
dependence. The second lSV contains the deviation from the average difference maps
and its temporal variation is described in the second rSV. In general, the number of
significant vectors is equal to the number of kinetic processes involved in a reaction.

2.2.16 Kinetic analysis of Laue data

A major step to perform the kinetic analysis involves the determination of relaxation
times of the associated kinetic processes. The time-traces of significant rSVs are a linear
combination of the true time-dependent concentrations of the associated intermediates.
The significant rSVs are fitted globally with a sum of exponentials. The equation for
fitting of \( j \) rSVs with \( i \) relaxation times can be written as

\[
   rSV_j = A_{0,j} + \sum_{i,j} A_{i,j} e^{-\frac{1}{\tau_i}}
\]

where \( j \) is the number of significant rSV, \( i \) is the number of exponentials required for the
fitting, \( A_{ij} \) is the amplitude of \( i^{th} \) exponent term for the \( j^{th} \) rSV. Associated relaxation
times (\( \tau_i \)) and corresponding macroscopic rate coefficients (\( \Lambda_i = 1/\tau_i \)) can then be
calculated which infer the number of kinetic processes evolved after reaction initiation.

2.2.17 Posterior analysis

In a time-resolved diffraction experiment, each time-dependent difference electron
density map \( \Delta \rho_i \) in a time-series may consist of a mixture of two or more intermediates.
The extent of this mixture depends on the underlying chemical kinetic mechanism. Once
the structures of these intermediates are known, the observed \( \Delta \rho_i \) can be fitted by the
calculated time-dependent difference electron densities using posterior analysis. This analysis is equivalent to kinetic target analysis (van Stokkum et al., 2004) and is performed using the program GetMech (Schmidt et al., 2004, Schmidt, 2008). A given reaction can be appropriately defined by several kinetic mechanisms; however for the analysis, a simplified mechanism should be selected. For example, two kinetic mechanisms that can be employed for the slower part of the PYP photocycle are shown in Fig. 2.18. In the first one (Fig. 2.18A), the intermediate pB forms from a not nearer specified source S with the rate coefficient k₁. The pB then decays to the ground state pG with k₂. In Fig. 2.18B, two intermediates pR and pB are involved. pB forms from the pR with the k₁ and decays into pG with the k₂. An alternate path from the pR to pG is also possible with k₃. All the calculations regarding posterior analysis are performed by Prof. M. Schmidt. However, for a better understanding, the results obtained from these analyses are also discussed in the thesis.

2.3 Time-resolved spectroscopy

2.3.1 UV-Vis radiation
The chromophore (Fig. 2.19) is responsible for the color of biomolecules. If this molecule is excited, the energy difference between the ground and excited state may fall within the range of either the visible or the ultraviolet spectrum. Therefore, these portions of the electromagnetic spectrum are appropriate for UV/Vis spectroscopic investigations.

### 2.3.2 Chromophore

![Figure 2.19 Chromophore of photoactive yellow protein (PYP). S denotes the thioester bond through which the chromophore is linked to the protein.](image)

The chromophore is an aromatic system in which mesomerically coupled double bonds exist. The valence band electrons are delocalized and can freely move within the bonding network. When a chromophore molecule is illuminated with visible light of a certain wavelength, configurational (trans/cis) changes occur. This may trigger conformational changes of the protein. For example, if the chromophore of photoactive yellow protein (PYP) absorbs a blue photon, it isomerizes from the trans to the cis configuration. As a result, large conformational changes may occur in the surrounding protein matrix of PYP.

### 2.3.3 Beer-Lambert law

When visible light passes through a sample, certain wavelengths are absorbed and others are transmitted. Absorption depends on the chemical species present in the sample. The absorbance of a chemical species in a given sample can be calculated using the Beer-Lambert law which is given by
where $A$ is the absorbance of a particular specie, $I_0$ is the intensity of incident light, $I$ is the intensity of light transmitted through the sample, $\epsilon$ is the extinction coefficient ($\text{cm}^2 \text{mmol}^{-1}$) also known as absorptivity, $L$ is the optical path length, $c$ is the concentration of that specie in the sample. The absorbance is reliable below ~1.5 absorption units (a.u.), because nonlinear effects start to appear for values higher than that.

### 2.3.4 Time-resolved spectroscopic investigations

Similar to time-resolved crystallography (Chapter 1 & 2), a laser pulse of a particular wavelength is used to trigger a reaction in a protein. The subsequent spectroscopic changes are then probed with visible light at different time delays $\Delta t_i$ after the laser pulse. Instead of probing structural changes with X-rays, here we monitor spectroscopic
changes. With time-resolved spectroscopy, the evolution as well as the decay of several intermediates in the protein can be examined.

2.3.5  Sample preparation

2.3.5.1  Protein solution in a capillary

The concentration of a protein sample is optimized such that its maximum absorbance $A_{\text{max}}$ lies in the range of $\sim 1$ a.u.. A suitable buffer (2.7M ammonium sulfate, 50mM sodium phosphate, pH 7.0) is used for the dilution of the protein in order to adjust its concentration. The optimum concentration varies for different proteins and its typical value varies from 1 to 4 mg/mL. A drop of 200–800 nL protein solution is placed in a glass capillary of diameter 1mm (Fig. 2.20 A). A similar sized drop of dilution buffer (not shown) is also placed in the capillary so that the absorption spectrum collected for the buffer can be subtracted from those obtained for the protein. This eliminates the effects of the dilution buffer on the protein spectra. Three to five drops of a protein may be mounted in a single capillary so that several experiments may be performed with the same capillary. This avoids the mounting of a new sample each time when the sample in a particular drop is bleached or destroyed. Epoxy is used to properly seal the capillary in order to avoid drying out of the protein solution.

2.3.5.2  Protein crystals crushed between the cover slides

Protein crystals may be unusually optically thick. PYP for example has a molar absorption coefficient of 45500 cm$^2$ mmol$^{-1}$ at 446 nm. The monitoring light is significantly absorbed. If the PYP concentration in its crystalline form is $\sim$100 mM, a crystal of thickness 200 $\mu$m corresponds to $\sim$ 100 a.u.. These high values are impossible to measure. To obtain absorption spectra whose absorbance unit $\sim$1 a.u., a crystal of
thickness 2 µm is required. To fulfill this requirement, crystals are crushed to produce a micron thin layer of crystalline sample. This significantly reduces the optical path of the probing beam through the sample. For the sample preparation, a small drop of stabilization buffer is placed on the siliconized cover slide (Hampton Research) of diameter 12 mm (Fig. 2.20B & C). A protein crystal is then looped and immersed properly in the drop of buffer. The stabilization buffer is not only necessary to prevent the disintegration of protein crystals but also it keeps the crystal moist. The crystal is covered by another cover slide of the same size and then gently squeezed in between to achieve the required thickness. This assembly is quickly sealed with epoxy. If the sample is not sealed properly, crystals of ammonium sulfate used in the stabilization buffer start to form (Fig. 2.20D).

### 2.3.6 Sample mounting

A setup where we have mounted the sample is shown in the Fig. 2.21A. A capillary Cp containing the protein solution can be easily mounted on the goniometer such that its long axis is perpendicular to both the laser and the visible light (IVL) (Fig. 2.21A & B). This ensures that the sample volume which is illuminated by the laser pulse can be probed. On the other hand, mounting of the cover slide is not that straightforward, because the cover slide has a flat circular area and cannot be placed perpendicular to both laser and visible light. To probe the illuminated volume, the cover slide has to be tilted by 40~45° towards the direction of laser light as shown in the Fig. 2.21C. The illuminated volume is further maximized by tilting the laser beam about ~30° from the horizontal direction (Fig. 2.21C). With these geometrical adjustments, a given sample volume that lies in a tilted
plane can be easily illuminated with the laser pulses. The illuminated volume can then be probed with the visible light.

![Image of the sample environment of our micro-spectrophotometer setup](image)

**Figure 2.21** A. Image of the sample environment of our micro-spectrophotometer setup. LP: laser pulse, IVL: incident visible light, TVL: light transmitted through the sample, Cm: camera, Cr: cryojet stream, Cp: Capillary. Cp containing protein solution is mounted on the goniostat. B. A drop of protein solution, S (orange) mounted in the capillary (light grey rectangle). The capillary is in the horizontal plane (dark grey). It is also perpendicular to both the laser (blue) and the visible light (green). C. A crystal, S (orange) is crushed between two cover slides (light grey ellipses). The cover slides are in a plane that is tilted (dark grey) towards the laser light. To optimize the overlap between laser and visible light, the laser beam is also tilted around 25~30° towards the visible light as shown.

### 2.3.7 Setup of a self-designed fast micro-spectrophotometer

A scheme for all the components involved in the setup is shown in Fig. 2.22A. For a reaction initiation in the sample, a nanosecond laser pulse (PL) from an Opotek HE 355II laser is used. A polarizer, POL (Glan Laser ThorLabs) is used to attenuate the laser beam as necessary. The beam is focused and coupled to a 600 μm fiber OF1, which is used to transport the light to the sample. The intensity of a laser beam should be tuned such that it contains enough photons to effectively trigger a reaction but does not destroy the protein.
Figure 2.22 A. A scheme of the in-house microspectrophotometer designed in our lab. PL: Opollete HE II pulsed laser, VIS: visible light source, SH1-2: Shutters, POL: polarizer, OF1-3: optical fibers, L1-8: lenses, M1-4: mirrors, G: grating, TC: cryojet. B. Scheme of the image intensifier tube equipped with a microchannel plate, MCP. A photocathode (in blue) emits electrons upon incident photons of a particular wavelength. $V_A$: voltage between input and output window, $V_B$: voltage across the MCP. The MCP comprises many channels (slightly tilted horizontal lines). The accelerated electrons generate secondary electrons (curved lines), which are converted to photons by the phosphor screen (in green). The amplified intensity is finally detected by CCD sensor (vertical line). C. A typical spectrum of the Xenon light source employed for optical monitoring of a protein sample.
Shutter, SH2 (ThorLab, SC10) is used to avoid accidental exposure to the laser beam. To probe the spectral changes in the sample after the laser excitation, a continuous visible light source (VIS, Asahi spectra Xe 300W) is employed. Its spectrum acquired from our spectrophotometer is shown in the Fig. 2.22C. Another shutter, SH1 (ThorLabs, SC10) is used to expose the sample to the visible light only for a short period of time. The minimum opening time depends not only on the shutter mechanism but also on the employed electronics. For our setup, a minimum opening time of 12~15 ms is possible. The visible light is coupled into tapered fiber, OF2, and focused to a spot size of 220 x 220 µm$^2$ at the sample position, S. Light transmitted through the sample is collected into a 400 µm optical fiber, OF3, which is focused using a lens (L6-L8) and mirror (M1) assembly (courtesy Bio-cars, enclosed in black dashed rectangular box as shown in the Fig. 2.22A). The light is then sent to the spectrophotometer, Shamrock 303 (Andor technology), where it is dispersed into different wavelengths by a diffraction grating, G. The dispersed beam is captured by a charge coupled device (CCD) sensor which is open for a certain period of time referred to as exposure time$^1$. In our spectrophotometer, the CCD camera (Andor iStar) is equipped with an image intensifier tube (Fig. 2.22B). The tube contains a photocathode which emits electrons while photons are incident on it. Depending on the voltage $V_B$, these photoelectrons reach the microchannel plate, MCP, of thickness 1 mm (Andor iStar manual). This plate consists of many channels whose widths are 10 µm. Since a high voltage $V_A$ in the range of 600 ~ 1000 V is applied across

---

$^1$ Words in italic in chapter 3 represent the terms that are related with synchronization and timing scheme of the micro-spectrophotometer.
the MCP plate, photoelectrons cascade down the channel. As a result, secondary electrons are generated. The clouds of electrons are then converted to visible light photons by a phosphor screen. This way, the incident light is amplified and reaches the CCD chip. The amplification factor depends on the voltage $V_A$, which is denoted as the MCP gain. A higher MCP value may be selected for those spectra having a weak photon count. Signal to noise ratio can also be improved by averaging multiple spectra. The number of spectra to be averaged is referred to as accumulation number. The image intensifier can also be time-gated such that the intensifier is open for a given period of time at a desired time delay $\Delta t_i$ after a trigger. The Q-switch pulse from the laser may act as a trigger. The duration for which the intensifier tube opens is referred to as pulse width. In a typical spectrophotometer, there is a minimum response time between receiving an external trigger as an input and the execution of a program initiated by the trigger. This response time is denoted as insertion delay here. This delay defaults to a certain value which varies for different spectrophotometers. The delay for our spectrophotometer is 34 ns. A cryojet stream (Oxford HT II) is used to control the temperature at the sample position. Temperatures in the range of -150 to 300 °C can be maintained by using the gas stream.

2.3.8 Timing scheme for synchronization

The synchronization scheme as shown in the Fig. 2.23 is described in the following. The laser flash lamp is operated with a frequency of 20 Hz. An electronic pulse of the laser flash lamp output is fed into a digital delay generator (DG645, Stanford Research Systems). As a result, the DG645 is externally synchronized with the laser flash lamp. A
100 µs pulse is generated with the DG645 and then delayed by 146 µs with respect to the flash lamp pulse. This pulse is used to externally trigger and activate the Q-switch. A delay of 146 µs is selected because with this delay the maximum intensity of the output laser pulse can be obtained. After 40 ns of the Q-switch trigger, the light pulse exits from the laser and subsequently excites the sample. The Q-switch output is sent to the iStar camera of the spectrophotometer which has an insertion delay of 34 ns as mentioned earlier. The value of insertion delay smaller than 40 ns ensures that the data collection sequence can be started 6 ns before the sample is hit by the laser pulse. All the important variable settings such as time delays, accumulation numbers, pulse widths, CCD exposure times, relaxation times between two successive frames and MCP gains are provided by the data collection software. These variables are explained in section 3.3.8. The electronic output pulse from the spectrophotometer also called gate monitor pulse is amplified with another digital delay generator (DG545 courtesy BioCARS). The amplified pulse is then delayed before it is used to operate the visible light shutter SH1 as shown in the Fig. 2.22A. With this generator, the time at which the shutter is opened can be exactly controlled and changed accordingly for all the time delays.

The timing scheme for synchronization has recently been modified in our lab (Purwar et al., 2013) (not used in this thesis). In the updated scheme, the pulse that activates the iStar CCD and collects a time-point is externally delayed by the DG645. The dependency of the exposure time on the time delays is eliminated. Consequently, the effects of dark currents that were large in the previous scheme can be greatly minimized. The flashlamp TTL output is directed to DG645 as an external trigger as in the original scheme, however the Q-switch is activated by a pulse from DG645 delayed by about 50 ms. This
provides enough time to open the Xe light shutter (SH1 in Fig. 2.22A). As a result, the first laser pulse can be utilized for the data collection purposes. Time-resolution is also improved.

2.3.9 Display of selected pulses during data collection

![Electronic pulses monitored with the oscilloscope during time-resolved spectroscopic experiments. Pulse 1 (yellow): reference pulse from the DG645; Pulse 2 (magenta): output pulse from the Q-switch of the laser; Pulse 3 (cyan): output pulse from the visible light shutter; Pulse 4 (green): output gate monitor pulse from the Andor iStar spectrophotometer. The 4th pulse denotes the time during which the image intensifier of iStar is open. A. A reaction is started with the laser pulse and changes are probed for 60 µs at the time delay of 770 µs. The minimum possible time 13 ms is shown during which the visible light shutter is open. B. An enlarged view of panel A. The delay of 146 µs between the flash lamp pulse (yellow) and the Q-switch pulse (magenta) is selected to maximize the intensity of the laser pulse. 3rd and 4th pulses are the same as described above.](image)

A digital oscilloscope (LeCroy40S, 400MHz) is used to monitor input and output pulses before and during the experiments. Four pulses can be inspected at the same time (Fig.
The selection of these pulses is described in the following. The reference pulse from DG645 that trigger and operate the Q-switch is selected as the first one to be displayed. This pulse is also used as a trigger for the oscilloscope. Output pulses from the Q-switch and from the visible light shutter are displayed as second and third ones, respectively. The last pulse to be monitored is the output gate monitor pulse from the Andor iStar spectrophotometer.

2.3.10 Description of the Andor Basic data collection software

The data collection software is written in an Andor proprietary programming language called Andor Basic. A list of desired time delays is generated with a Matlab script such that all the time delays are equally distributed on a logarithmic time scale with at least 3 to 4 time points per logarithmic decade. Typically a time range from several nanoseconds to seconds can be covered. The Matlab generated time delays are scripted manually into the data collection sequence. The correct exposure times has to be determined for each time delay. The exposure time must at least be equal to the sum of a particular time delay and the corresponding pulse width. This is one of the limitations for the exposure time mandated by the Andor spectrophotometer. Another imperative limitation is that the minimum possible exposure time is 2 ms. With our current setup, time points as fast as 20 µs can be investigated. For time points faster than that, not enough photons reach the sensor. For fast time delays a pulse width of 20 µs is selected to obtain absorption spectra. Then the corresponding exposure times, that fulfill the minimum criteria set by the instrument as described before, are computed. For time scales such as 25 µs ~ 500 ms, the ratio of the pulse width and the exposure time falls between 0.8 ~ 0.04 (Fig. 2.25A). The effects of the dark current are negligible for these spectra (Fig. 2.25D). In order to further
enhance the intensity of these spectra, a MCP gain of 50 is applied. With this gain, smooth spectra can be obtained for time delays from 25 µs to ~ 500 ms. In the case of longer time delays (1 ~ 5 s), the exposure times become relatively large due to instrument limitations; however, the pulse width is short as mentioned earlier. As a consequence, the ratio of pulse width to exposure time is extensively small (Fig. 2.25B) and the spectra become very noisy (compare Fig. 2.25D & E). In order to reduce the effects of dark current, a relatively large pulse width between 1~2 ms can be selected. This leads to a high pulse width to the exposure time ratio (Fig. 2.25C), which minimizes the influence
of the dark current (Fig. 2.25F). If the pulse widths are long and a MCP gain of 50 is applied, the spectra start to saturate. The pulse width must be selected such that both the dark current and saturation effects are the smallest possible. For the longer time delays (1 ~ 5 s), pulse widths from 170 to 200 µs are suitable. A common MCP gain of 50 is chosen for all the time delays to minimize the background effects. These background noises originate when different MCP gains are selected for different time points. To improve the signal to noise ratio even more, multiple frames can be averaged for each time delay. About forty to fifty accumulations are necessary for time delays up to 100 µs whereas 15 ~ 20 exposures are enough for longer time delays. Once all the parameters

Figure 2.25 A. A charge couple device (CCD) of the Andor iStar with full vertical binning (FVB) as a read out mode. 1024 pixels are distributed on the X-axis. These pixels can be mapped to a given wavelength range. Absorption spectra of PYP plotted as a function of B. CCD pixels before calibration. C. wavelength after calibration.
such as pulse widths, exposure times, accumulations and MCP gains are set for each time delay, transmitted spectra at these settings can be collected. First, the background transmission spectra $T(\lambda)_{\Delta t}^{bg}$ without any visible light and then the reference spectra $T(\lambda)_{\Delta t}^{ref}$ with visible light and without any protein are collected for each time delay. After that, the signal spectra $T(\lambda)_{\Delta t}^{sig}$ at the sample position are collected. From these, absorption spectra at different time delays $A(\lambda)_{\Delta t} = -\log \left( \frac{T(\lambda)_{\Delta t}^{sig} - T(\lambda)_{\Delta t}^{bg}}{T(\lambda)_{\Delta t}^{ref} - T(\lambda)_{\Delta t}^{bg}} \right)$ can be calculated. As a reference, dark absorption spectra $A(\lambda)_D$ without any laser excitation are also obtained before and after collecting the time-resolved data. The data collected at different time delays are saved according to their progression in time and a comprehensive time series is obtained.

During the data collection, wavelengths in the absorption spectra (Fig. 2.25B) are by default mapped to pixel positions on the CCD chip (Fig. 2.25A). There are total 1024 pixels in the x-direction. The pixel positions must be transformed into wavelengths (Fig. 2.26A & C). This requires calibration. Our system can be easily calibrated by using a known calibration source such as CAL2000. It produces mercury and argon atomic emission lines from 253 to 922 nm, whose pixel positions are known. Few distinct peaks are selected and their pixel positions are captured. These data points can be fitted by a linear, quadratic or a cubic fit. Although the spectral resolution of the detector is 12 nm due to its point spread function, the use of grating with 1199 lines/mm enables us to distinguish peaks that are only 0.1 nm apart with an accuracy of ±0.2 nm.

2.3.11 The Matlab data analysis program
For the analysis of the dark and the light data, a self-written Matlab program is used. The absorption spectra for each time delay are plotted (Fig. 2.27A). There is a slight offset relative to the X-axis which is different for each spectrum. These offsets are shifted to the offset of the corresponding dark spectrum (Fig. 2.27B). For that, the spectral range from 550 to 700 nm, where no spectral changes occur, is used. Once brought to the same offset, time-dependent difference absorption spectra are calculated. The time and wavelength dependence of the difference spectra are analyzed by SVD. This analysis has already been discussed for time-resolved crystallography, where it is applied to the difference electron density maps (Sec. 2.2.15). The basic concept is the same whether SVD is applied to the X-ray data or to the spectroscopic data. Here, we focus on the details that apply to the spectra. A wavelength range in which largest changes in the difference absorption spectra occur is selected. For example, in Fig. 2.27 significant changes are observed in the wavelength range from 400

**Figure 2.26** Typical examples of time-resolved absorption spectra collected on PYP solution at ambient temperature. A. A time series of ten absorption spectra from 30 µs to 2 sec. As evident, offsets corresponding to each spectrum are slightly different from each other. B. The absorption spectra are shifted with respect to the reference spectra and bought to the same level.
to 520 nm and therefore this range should be selected for the SVD analysis. Spectral changes in other than the specified regions are omitted from the analysis because they do not contain signal and add only noise. Difference spectra in the selected wavelength range are arranged in a data matrix for the SVD (Fig. 2.17). After decomposition, the obtained left (lSVs) and right (rSVs) singular vectors contain spectral and kinetic information, respectively. All the significant rSVs are globally fit with a sum of exponentials using a fitting routine written in Matlab. Data fitting is tested with different number of exponentials. Too few exponentials do not fit the data properly, whereas too many exponentials do not improve the quality of fitting any further. Associated relaxation times ($\tau_i$) and corresponding rate coefficients ($\Lambda_i = 1/\tau_i$) can then be calculated. $\tau_i$ infer the number of kinetic processes that evolve after a reaction is initiated.

## 2.4 Biochemical procedures

So far, we have described the employed techniques with experimental setups and associated hardware. In the following sections, we describe how biological samples such as catalase and PYP are prepared for our experiments. From the purification to the crystallization of these proteins, all necessary steps are described.

### 2.4.1 Beef liver catalase (BLC)

#### 2.4.1.1 Purification

Bovine liver catalase (BLC) is obtained from sigma (C3155). To purify the protein, size exclusion chromatography is used in which proteins are separated based on their sizes or their molecular weights. Superdex 200 (GE) is selected as a gel filtration medium which is equilibrated with buffer consisting of 10mM Tris, 50mM NaCl, pH 8.5. Protein
molecules larger than the gel pore size are unable to diffuse into the gel and therefore eluted first. On the other hand, smaller protein molecules penetrate into the gel pores and elute later. The eluted catalase is collected in small fractions of 2 mL. Concentration of all these fractions can be calculated using the Beer-Lambert law. UV-Vis spectra obtained for each fraction provides the absorbance at 405nm ($A_{405}$). The extinction coefficient ($\varepsilon$) at the same wavelength is also known (Reid et al., 1981). Its value is 420000 cm$^2$mmol$^{-1}$ for the tetramer. This allows us to quantify the catalase.

**2.4.1.2 Purity assessment using absorption spectra**

The 405nm band of catalase is due to the presence of four heme prosthetic groups which absorb at 405 nm. Aromatic residues are present in almost all proteins, and they absorb at 276 nm. The ratio ($A_{405nm}/A_{276nm}$) of the absorbance at 405 nm to that at 276 nm can be used to assess the purity of catalase. Only those catalase fractions whose ratio is larger than one are pooled. These protein fractions are concentrated to ~15 mg/mL. To the concentrated and purified catalase solution, 1-5 wt % of NH$_4$OH (30% in water) is added right before starting crystallization trials.

**2.4.1.3 Crystallization**

Catalase crystals are grown by the hanging drop vapor diffusion method. The crystallization conditions are used as reported previously (Ko et al., 1999). Briefly, protein concentration of 12~13 mg/mL containing NH$_4$OH as described before is used. Four microliters of the catalase solution is mixed with an equal volume of the reservoir solution that contains 45~60 mM magnesium formate. After 2-3 weeks, crystals of typical size 100 x 100 x 200 µm$^3$ are obtained.

**2.4.2 Photoactive yellow protein (PYP)**
2.4.2.1 Overexpression of PYP in E. coli

PYP-DNA from the original bacterium E. halophila is expressed in the host bacterium E.coli. Luria broth with ampicillin is used a medium to grow bacteria. Once the OD$_{600}$ reaches to ~0.6, the bacterial growth is stopped by keeping them on the ice. Isopropyl β-D-1-thiogalactopyranoside (IPTG) is added to a final concentration of 1mM to induce overexpression of apo-PYP. Bacteria are further grown at 16 °C for ~22 hours. These bacteria are then collected by centrifugation and cell walls are ruptured by sonication. Lysozyme and protease inhibitor are added into the culture right before sonication. P-coumaric anhydride (pCA anhydride; 3M DCC, p-coumaric acid solution, DMF as solvent) is added to apo-PYP. The holo-PYP attached with the Chromophore is further purified to remove unwanted proteins or impurities.

2.4.2.2 Purification

Histidine-tags (His-tags) introduced in the PYP sequence were used to purify the PYP. For this, an immobilized metal ion affinity chromatographic column (IMAC) charged with Ni$^{2+}$ resin is used. When PYP solution passes through the column, His-tag PYP are adsorbed to the resin and remain in the column. These His-tag proteins are then displaced by imidazole and are eluted from the column. In order to cleave the His-tags from the PYP, Enterokinase is used. For further purification of PYP, ion exchange chromatography where negatively charged PYP are adsorbed to positively charged gel medium (Q-sepharose GE healthcare) is used. A gradient from zero to 1M NaCl is used to elute PYP depending on its charge. Fast protein liquid chromatography (FPLC, AKTA prime) is used to operate all chromatographic columns involved in the purification process. For long term storage, the purified PYP is flash-frozen in liquid nitrogen.
Glycerol (5%) is added to the purified protein because it prevents ice formations and promotes stability of the protein. All the purification steps are assessed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). For this electrophoresis, proteins samples are denatured by heating and stabilized by addition of an anionic detergent SDS and a reducing agent β-mercaptoethanol. Denatured samples pass through the gel. Upon applying an electric field across the gel, linearized proteins start to move according to their sizes. To visualize the separated bands corresponding to different protein samples, staining and destaining are performed using conventional dyes (Lameelie blue dye).

2.4.2.3 Crystallization

Large crystals of PYP were grown as reported in (Borgstahl et al., 1995). Briefly, crystals are grown with sitting drop vapor diffusion method using a stabilization buffer containing 2.7M ammonium sulfate, 50mM sodium phosphate, pH 7.0. PYP of concentration 18–22 mg/mL is used for the crystallization trials. Crystals grow very slowly and they are highly twinned. To obtain high quality crystals and to accelerate their growth, these twinned crystals are used as a seed. To prepare a microseed solution, a few crystals are crushed in the stabilization buffer. The micro-seed is added to the protein drops after equilibration at ambient temperature for 2 to 3 days. To add the micro-seeds, a cat whisker is used. Large crystals with typical sizes of ~120 x 120 x 700 µm³ are then obtained.

2.5 Experimental details to investigate interaction of nitric oxide (NO) with catalase

Two techniques, monochromatic X-ray crystallography and micro-spectrophotometry are used to investigate the interaction of NO with catalase. In the following sections, we will
first describe the details of crystallographic experiments and then those of spectroscopic experiments. The latter experiments are not performed with our newly designed in-house spectrophotometer but with a much slower spectrometer available at BioCars.

2.5.1 Monochromatic X-ray crystallographic experiments

2.5.1.1 Preparation of different forms of Catalase

Three crystal forms are prepared (Fig. 2.28). Crystals obtained directly from the well plate are termed as Cat-NH$_3$ because NH$_4$OH was added into catalase solution as mentioned in the Sec. 2.4.1.2. Consequently, NH$_3$ binds to the active site. To prepare a second form where the NH$_3$ is removed, Cat-NH$_3$ crystals are first soaked overnight in 80 mM magnesium formate solution, pH 6.7 and then washed two more times with the same solution. These crystals are referred as Cat-5 (Fig. 2.28). Since the goal here is to bind nitric oxide (NO) to the Cat-5 crystals, the third form of the crystals is prepared by soaking the Cat-5 crystals with an appropriate NO generator. For this, the crystals are soaked for 5 minutes in a solution containing 80 mM magnesium formate, 150mM Bis-Tris as a buffer and 1-(N,N-diethylamino)diazene-1-ium-1,2-diolate (DEANO, Fig. 2.29A) as a NO generator. The concentration of DEANO was varied from 10 to 200 mM.
The reaction of NO with O₂ generates acidic products (Cotton et al., 1999). This can cause a pH shift, which can be avoided by immersing these crystals in a well buffered solution. To minimize the escape of NO into the atmosphere during the crystal nitrosylation process, a simple device is used. As shown in the Fig. 2.29B, this apparatus consists of two glass plates with the mating surfaces ground to provide a very tight seal.

The glass vessel V has a hemispherical depression with a capacity of ~ 200 µL. A Cat-5 crystal is placed in the depression, which is then filled with the stabilization buffer/DEANO mixture. The depression is quickly covered with the top lid L so that NO escape can be prevented. A heavy weight was placed on the top of upper plate to maintain high internal pressure. Thus obtained crystals are termed as Cat-NO (Fig. 2.28).

2.5.1.2 Data collection

All three crystal forms are soaked in a cryoprotectant consisting of 20% polyethylene glycol 4000, 20% sucrose, 50mM Tris, pH 8.5. For the Cat-NO crystals, it is imperative to quickly remove the crystals from the pressure cell and immediately freeze them in the cryojet because NO starts to diffuse out of the crystals as soon as the pressure are
released. NO occupancy at the active site drops to an undetectable level within ~10 minutes unless the crystal is frozen. For all three crystal forms (Cat-NH₃, Cat-5 and Cat-NO), data sets are collected at cryogenic temperatures. Mosaicity of these crystal forms is estimated prior to the actual data collection. The crystals are rotated through 180° in steps of 0.5° to cover the reciprocal space for a complete data set.

2.5.1.3 Data analysis

Data sets for the crystal forms Cat-NH₃, Cat-5 and Cat-NO are reduced with Mosflm (Sec. 2.1.17). Reflection intensities are scaled and merged with SCALA. From these intensities, structure factor amplitudes are calculated using TRUNCATE. The three dimensional crystal structure of beef liver Catalase is already solved for the orthorhombic crystal form and available as a PDB entry 4BLC (Ko et al., 1999). Therefore, molecular replacement is not required in this case, and the model 4BLC is used directly as an initial model. To prepare a basic starting model, all the water molecules and the cofactors (nicotinamide adenine dinucleotide phosphate, reduced NADPH) are removed from the model. This model is refined using rigid body refinement at 3Å. After that, simulated annealing refinement using a 2000K protocol up to full resolution is performed. Positional conventional refinement is performed until convergence. B-factors are refined using a 40-step restrained B-factor refinement. Water molecules are searched up to 3σ level in the difference electron density map (F_{obs} - F_{calc}) and added to the structure. Some of the waters are found in cavities. These cavities are determined with the program ‘voidoo’ (Kleywegt & Jones, 1994) using a probe with radius 1.4 Å on a 0.5 Å grid.
2.5.1.4 **Insertion of small molecules such as NH\textsubscript{3} or NO**

The active site heme and its vicinity are examined for the presence of any putative molecule. If a positive feature is identified in the difference map (F\text{obs} - F\text{calc}), it can be interpreted as NH\textsubscript{3} or NO, depending on how the crystal was treated. The occupancy of the ligand is roughly estimated by calculating the total number of electrons in the positive feature. For that, the positive electron density feature is integrated using the program *Probe* (Srajer *et al.*, 2001). The electron count is less than 10 for Cat-NH\textsubscript{3} form and less than 15 for Cat-NO form. Therefore, it is necessary to determine the exact occupancy of the ligand in each crystal form. For that, the nitrogen atom of either ligands, NH\textsubscript{3} or NO, is connected to the heme iron at a fixed distance using a very weak bond length restraint. NO is also aligned into a particular orientation with weak angular restraints which is not applicable for the Cat-NH\textsubscript{3}. Refinement commences using the grouped occupational refinement option in CNS. Then, positional conventional refinement is performed to refine the position and orientation of these ligands.

2.5.1.5 **B-factor averaging**

For the Cat-NO, B-factors are determined for the nitrogen and the oxygen atoms of the NO. B-factors for these two atoms as well as for the iron and the four nitrogen atoms of the heme are averaged (Fig. 2.30B). The average B-factor is then assigned to both N and O. Similarly, an average B-factor is obtained for NH\textsubscript{3}. The obtained B-factors are refined with restrained B-factor refinement and then conventional positional refinement is performed. Both of these refinements are repeated until convergence is reached.
2.5.1.6 Geometry of the ligand molecule

To describe the orientation of the ligand molecule, the tilt (τ) angle, the bend angle (φ) and their sum (α) with respect to a given plane can be used (Fig. 2.30). Two ligands NH₃ and NO are bound to the catalase. The ligand NH₃ consists of a single atom (ignoring hydrogen atoms). For NH₃, only the tilt angle with respect to normal of the heme plane can be determined where heme plane is defined by the four porphyrin nitrogens (Fig. 2.30A). For NO, all angles τ, φ and α with respect to the heme plane can be calculated (Fig. 2.30B). The iron out of plane distance can be calculated using Planefit (Tab. 3.1) (Nienhaus et al., 2005).

2.5.2 Micro-spectrophotometric experiments (performed at BioCARS)

2.5.2.1 Data collection and analysis

A microspectrophotometer (4DX) equipped with a Deuterium/Tungsten light source (DH2000, Ocean Optics) and a USB mini-spectrometer (USB 2000, Ocean Optics) is mounted in line with the X-ray beam at BioCARS 14BM-C. The spectroscopic
experiments for the three crystal forms, Cat-NH$_3$, Cat-5 and Cat-NO are performed at 100 K. The absorption spectra, $A(\lambda)_{\Delta t}$ of all species are recorded for up to 10 minutes in intervals of 10 s during X-ray exposure. Subsequent exposure times can be written as 10s, 20s, 30s…..600s. Difference spectra, $\Delta A(\lambda)_{\Delta t}$ are generated by subtracting the $A(\lambda)_{\Delta t}$ from the reference spectrum, $A(\lambda)$ recorded before X-ray exposure. Absolute differences are integrated from wavelength 500 to 700 nm for each exposure time. They are denoted as $\Delta A_{\text{int}}(t)$ and plotted against exposure time. Data are fitted empirically either by a single exponential, a sum of two exponentials or by an exponential and a linear phase.

2.6 Experimental details to investigate the kinetic dose limit in PYP

2.6.1 Data collection

Two PYP crystals of dimensions 170 x 170 x 700 $\mu$m$^3$ (PYP1) and 150 x 150 x 900 $\mu$m$^3$ (PYP2) are equilibrated at pH 7 and mounted individually in a capillary as described in sec. 2.2.9. The crystal PYP1 is used for actual experiment and PYP2 for a control experiment (Schmidt et al., 2012). The time-resolved technique is described extensively in sec. 2.2.10 and all the variables are defined in Sec. 2.2.11. Here we report some details specific to this experiment. Four X-rays exposures ($N_p = 4$) are accumulated per diffraction pattern. The slower part of the PYP photocycle (Fig. 1.4) is selected to investigate radiation damage. A time range from 256 $\mu$s to 32 ms is used. A short time-series, containing $N_t = 8$ time delays (256 $\mu$s, 512 $\mu$s, 1 ms, 2 ms, 4 ms, 8 ms, 16 ms and 32 ms) and one dark exposure, is obtained. $N_o = 20$ different crystal orientations are used. Between subsequent orientations, the crystal is translated by 22 $\mu$m and therefore
total crystal translation is 440 µm (Fig. 2.31B & Fig. 2.14B). The entire experiment is repeated 12 times and corresponding 12 short time-series are obtained (Fig. 3.3A). As a control experiment, crystal PYP2 is used to examine the damage due to laser pulses. The same protocol as described above is used without exposure to X-rays. The experiment is repeated 16 times which are equivalent to 16 virtual time-series. During these 16 time-series, 5 dark exposures are collected at equal intervals. The crystal is exposed to 512 laser pulses per orientation. A total of 10240 laser pulses are accumulated during the entire control experiment.

2.6.2 Absorbed dose calculation

![Diagram](https://example.com/diagram.png)

**Figure 2.30** A model used to determine the common area $F_c$. X-ray beam falls on the PYP crystal. As the crystal is re-oriented by $\Delta \theta$, it is irradiated from one direction (red rectangle) to another direction (yellow rectangle). **A.** The common area (colored orange) is $F_c$ which is used to calculate the common volume $V_c$ (not shown). **B.** A sequence of angular settings, $\Delta \theta$, (in degrees). These settings are also separated by translations of 22 µm. Five settings are fully or partially exposed with the employed X-ray beam. The orange bars denote the relative sizes of the $V_c$ for each angular setting. Figure from Schmidt et al., 2012.

Since the goal is to find out how many data sets can safely be collected without extensive damage, the energy deposited in the crystal due to the ionizing radiation must be determined. To calculate the absorbed dose in Joule/kg = Gy, the program ‘Raddose’ (Murray *et al.*, 2004) is used. If the crystal PYP1 is irradiated by a single X-ray pulse of
90 (h) \times 60 (v) \, \mu m^2$, the illuminated voxel is equal to beam size times the crystal thickness ($t = 170 \, \mu m$). The dose absorbed by the illuminated voxel, $d_{vox}$, is calculated. If $l_{eff}$ is length of the crystal that is exposed to X-ray beam, the total irradiated crystal volume is $l_{eff} \times v \times t \, \mu m^3$. The sum of overall crystal translation in 20 orientations (440 \, \mu m) and X-ray beam size in the horizontal direction (90\mu m) provides a $l_{eff}$ of 530 \, \mu m (Fig. 2.14B). The dose absorbed by this crystal volume ($d_{crys}$) due to each single X-ray pulse is computed as

$$d_{crys} = \frac{h \times v \times t}{l_{eff} \times v \times t} \, d_{vox} = \frac{h}{l_{eff}} \, d_{vox} \quad (2.28)$$

Each short time-series consists of 8 time delays, one dark exposure and one exposure due to edge scan (Sec. 2.2.10). If 20 crystal orientations and 4 X-ray pulses per diffraction are taken into account, the average dose absorbed in one comprehensive time series, $d_{ts}$, can be written as

$$d_{ts} = d_{crys} \times (N_t + 2) \times N_o \times N_p \quad (2.29)$$

where $N_t$, $N_o$, and $N_p$ are defined in Sec. 2.2.11. In order to determine the total dose ($D_{nc}$) during the entire experiment, the dose for one time-series is multiplied by the total number of collected time-series. The intensity values are used to address radiation damage. The steps involved in the determination of intensities from the raw Laue diffraction pattern are described in Sec. 2.2.12-13 & Fig. 2.15A. $\langle I \rangle$ and $\langle I/\sigma_I \rangle$ are plotted as a function of uncorrected dose $D_{nc}$ (Fig.3.3A). Two types of correction are required. The first correction is due to the subsequent crystal settings. Since the thickness of the crystal ($t$) is larger than the vertical X-ray beam size ($v$), each time when the crystal is rotated by an angle, $\Delta \theta$, from one orientation (in red) to another (in yellow), some fresh
crystal volume is exposed (Fig. 2.31A). As a result, the dose deposited into the given crystal volume is decreased. To calculate the correction term, the common area $F_c$ and the common volume $V_c$ between two consecutive crystal settings can be determined (Fig. 2.31). If the size of X-ray beam $h = 90 \, \mu m$ is divided by the translation step in each setting (22$\mu m$), five crystal settings are affected as shown in Fig. 2.31B. Among these 5 settings, three are covered fully by the X-ray beam while two are covered partially. The $V_c$ for each of the five neighboring settings are calculated and averaged (Fig. 2.31B). The average volume, $\langle V_c \rangle$ is the first correction term for the absorbed X-ray dose for all the short time-series (Fig. 2.31B). The second type of correction is required due to the edge scan method. Edge scan is performed once at the beginning of each crystal setting. After each edge scan, the x and y coordinates of the goniometer are recorded. The relative vertical displacements, $y'$, of the crystal across the X-ray beam for each setting can be calculated (Fig. 2.31). The individual $y'$ are averaged to obtain $\langle y' \rangle$. The crystal displacement about $\langle y' \rangle$ exposes a fresh crystal volume to the beam that further reduces the deposited dose. Using $\langle y' \rangle$, the dose can be corrected for each sweep. In addition, the crystal damage due to laser pulses during the entire experiment must be determined and corrected. Since the crystal is exposed to only laser pulses without exposing to the X-ray pulses, these laser pulses can be assumed equivalent to virtual X-ray doses. Once the virtual dose is calculated, corrections due to different crystal orientations and due to the edge scan can be applied to the virtual data as described for X-ray exposure. The adjusted virtual dose corresponding to each of the 16 virtual time-series $D_L$ is obtained. $\langle I \rangle$ and $\langle I/\sigma_I \rangle$ corresponding to all five dark data sets are plotted as a function of adjusted virtual dose, $D_L$ (Fig. 3.3B). The intensities and their standard deviations both are fitted with
straight lines. The y-intercept at zero absorbed-dose is normalized to unity and then the slope ($S_L$) is determined. Assuming that the damage due to the laser pulses and X-ray pulses are independent events, the intensities ($I$) corresponding to all 12 time series, for which X-ray as well as laser exposures are used, are corrected for the laser damage (Fig. 3.3B). The Intensities free from the laser damage ($I_{\text{free}}$) are given by

$$I_{\text{free}} = I(-S_LD_{\text{corr}} + 1)$$

(2.30)

### 2.6.3 SVD analysis

For all the 12 time-series, difference structure factor amplitudes are calculated as reported (Sec. 2.2.14)(Ren et al., 2001, Ihee et al., 2005). PDB (Berman et al., 2002) entry 2PHY is used to provide phases to calculate time-dependent difference maps. A total of 96 difference maps are obtained corresponding to 12 time-series. Each of these time-series is analyzed by SVD (Schmidt et al., 2003, Zhao & Schmidt, 2009). The difference electron density at the chromophore region is included in analysis (sec. 2.2.16). Grid points including the chromophore as well as the amino acid residues that line the chromophore pocket such as Tyr-42, Glu-46, Met-100 and Arg-52 are masked out (Schmidt et al., 2012). The mask is evolved by allowing only grid points that are above +2.5\(\sigma\) or below -2.5\(\sigma\) in at least one of the difference maps of the time-series. The SVD is performed on the masked difference maps and right singular vectors (rSVs) are obtained (Fig. 3.4). From the fitted rSVs, relaxation times are extracted. The kinetic mechanism shown in Fig. 2.18A is used for the Posterior analysis. If no rising phase is detected in the rSVs, the mechanism used for the analysis is given in the dashed box of
Fig. 2.18A. Once microscopic rate coefficients $k_i$ are obtained, their inverse relaxation times ($\tau_{post}$) can be calculated (Fig. 3.6A).

2.7 Experimental details to investigate the pH dependence of PYP photocycle

2.7.1 Data Collection

PYP crystals of size $\sim 100 \times 100 \times 700 \, \mu m^3$ are soaked for 5 minutes in stabilization buffer (2.7M Ammonium sulfate, 50mM Sodium phosphate) adjusted to pH 4, 7 and 9 (Tripathi et al., 2012). The crystals are mounted in capillaries (Sec. 2.2.9 & Fig. 2.14A) and equilibrated for at least 12 hours in the respective buffer before data collection. The diffraction experiments for all three pH conditions are performed at 25 °C. The slower part of the photocycle is selected for the time-resolved investigation. At pH 4 a time range from 4 µs to 1 sec is employed and distributed through 18 time points ($N_t$). At pH 7 eighteen time delays are collected in the range from 1 µs to 512 ms, whereas at pH 9 twenty time delays are collected from 1 µs to 1 s. A comprehensive time-series, consisting of $N_t$ time delays and one dark exposure without any laser illumination, is collected for each pH condition. To recover the dark state, the waiting time between two consecutive laser pulses is 30 s, 4 s and 12 s at pH 4, 7, and 9, respectively. Twenty different crystal orientations ($N_o=20$) are used. X-ray pulses ($N_p$) accumulated per diffraction pattern are 3, 11 and 9 in the case of pH 4, 7 and 9, respectively. The edge scan is used to maximize the overlap of laser and X-ray illuminated volume (Sec. 2.2.10 & Fig. 2.14B). The time-resolved Laue data sets are processed by Precognition/Epinorm as described in Sec. 2.2.12-13. Phases are determined from PDB entry 2PHY. Difference electron density maps on the absolute scale are calculated. The difference electron
density of the chromophore region is used for the SVD analysis. From rSVs, relaxation times ($\tau_1 \& \tau_2$) are calculated as described in Sec. 2.2.16.

### 2.7.2 Kinetic analysis

For the posterior analysis (Schmidt, 2008), the kinetic model shown in Fig. 2.18B is used. The rate coefficients, $k_i$ of the mechanism are refined against observed electron density maps and concentration profiles are determined. Once $k_i$’s are known, the corresponding relaxation times ($\tau_p$) can be calculated.

### 2.7.3 Extrapolated maps

The difference maps corresponding to the time points $t_i < \tau_1$ contain structural information about the pR intermediates, while those time points $\tau_1 < t_i < \tau_2$ contain information about pB. The maps in the respective pR and pB regions are averaged (Tripathi et al., 2012). Fourier inversion of these average maps provide time-independent difference structure factors ($\Delta F_{tind}$) for both pR and pB. Extrapolated structure factors $F_{tind}^{ext}$ are then calculated as given by

$$F_{tind}^{ext} = F_{d}^{calc} + N\Delta F_{tind}$$

where $F_{d}^{calc}$ is structure factor of the dark state model and N is the multiplication factor. N has to be determined such that it compensate the missing scale between $\Delta F_{tind}$ and $F_{d}^{calc}$. To do that, we start with N=1 and corresponding $F_{tind}^{ext}$ are calculated. From these $F_{tind}^{ext}$, extrapolated electron density maps are obtained. The negative electron density features in these extrapolated maps are integrated. The factor N is increased and the above steps are repeated. Once N reaches up to a certain characteristic $N_c$, the procedure is stopped because after $N_c$ large negative features start to appear in the extrapolated
maps. Using the $N_c$ and respective $F^\text{ext}_{\text{tind}}$, conventional extrapolated electron density maps are obtained (Schmidt et al., 2004, Tripathi et al., 2012). The structural models for pR and pB are interpreted into the extrapolated difference maps and their atomic models are refined against the $|F^\text{ext}_{\text{tind}}|$ using CNS (Sec. 2.1.18). The improved model phases ($\phi_{\text{new}}$) are combined with the $|\Delta F_{\text{tind}}|$ to calculate new extrapolated structure factors and also new extrapolated maps.

2.8 Experimental details to investigate the temperature dependence of the PYP photocycle

The effect of temperature on the structure and kinetics of the PYP photocycle is investigated using two techniques: time-resolved crystallography and time-resolved spectroscopy. In the first few sections, experimental details related with crystallography are described. Spectroscopic experimental details are described in the later sections.

2.8.1 Time-resolved crystallography

2.8.1.1 Data collection

PYP crystals of typical sizes 100 x 100 x 700 $\mu$m$^3$ are mounted in capillaries as described in Sec. 2.2.9 (Fig. 2.14A). Diffraction experiments are performed at 14 different temperatures. The temperature is varied from -40 °C to 70 °C. For experiments above 30°C, the capillaries must be insulated from the brass pins as shown in Fig. 2.32A. Otherwise severe distillation effects due to the cold brass pin dry out the crystals.
Figure 2.31 A. Protein sample mounted in the capillary for \( T > 30 \, ^\circ\text{C} \). \( P \): brass pin, \( I \): insulating glass capillary, \( E \): epoxy glue, \( c \): PYP crystal (colored yellow), \( S \): stabilizing solution. B. Laser illuminated volume of the crystal (red dashed box). Capillary wall (thick black line) is temperature controlled. Black arrows show the directions of heat diffusion. Figure from Schmidt et al., 2013.

For temperature up to 40 °C time delays range from 2 ns to several seconds whereas for temperatures equal to or higher than 50 °C, the time delays start at 100 ns (Tab. 2.1). The waiting time between two subsequent laser pulses is varied from 1 s at higher temperatures to 20 s at lower temperatures (Tab. 2.1). Twenty different crystal orientations (\( N_o \)) are used. For the data collection, edge scan is used as described before. Data collected at all temperatures are processed with Precognition/Epinorm (Sec. 2.2.12-13). Weighted difference structure factor amplitudes are calculated (Sec. 2.2.14). With these amplitudes and phases from the PDB entry 2phy, time-dependent difference maps \( (\Delta \rho^{\text{obs}}) \) are calculated.

2.8.1.2 Kinetic analysis

SVD analysis is performed using the difference electron density maps. Relaxation times ( \( \tau \) ) and their inverse relaxation rates are obtained for each of the 14 temperature settings from the rSVs as described in Sec.2.2.16. The observed relaxation rates are plotted as a function of temperature and fitted by the Van’t Hoff-Arrhenius
equation (Eq. 1.4). From the fitting, prefactor $\nu$ and activation energy $E_a$ are calculated for the relaxation rates $\Lambda_2, \Lambda_3 \& \Lambda_4$.

**Table 2.1** Data collection statistics of PYP at different temperatures. Table from Schmidt et al., 2013.

<table>
<thead>
<tr>
<th>$T_{jet}$ (°C)</th>
<th>$T_{diode}$ [K]</th>
<th>$N_p$ (time points)</th>
<th>time range</th>
<th>wait time</th>
<th>completeness (last shell)</th>
<th>$I/\sigma_I$ (last shell)</th>
<th>$R_{merge}$ e (%)</th>
<th>$R_{scale}$ (%) @Δt</th>
<th>$\Delta \rho_{max}/\sigma_{\rho_p}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-40</td>
<td>235.5</td>
<td>31</td>
<td>2 ns to 15 s</td>
<td>20 s</td>
<td>82.5 (75.9)</td>
<td>16.9 (11.0)</td>
<td>8.3</td>
<td>7.7 (3 µs)</td>
<td>-6/10 (3 µs)</td>
</tr>
<tr>
<td>-30</td>
<td>245.1</td>
<td>27</td>
<td>1 ns to 4 s</td>
<td>12 s</td>
<td>77.3 (73.4)</td>
<td>28.3 (21.8)</td>
<td>5.3</td>
<td>6.8 (2 µs)</td>
<td>-11/13 (2 µs)</td>
</tr>
<tr>
<td>-20</td>
<td>254.7</td>
<td>34</td>
<td>2 ns to 8 s</td>
<td>12 s</td>
<td>82.7 (78.2)</td>
<td>29.3 (21.4)</td>
<td>4.9</td>
<td>5.4 (4 µs)</td>
<td>-11/11 (4 µs)</td>
</tr>
<tr>
<td>-15</td>
<td>259.5</td>
<td>29</td>
<td>2 ns to 8 s</td>
<td>10 s</td>
<td>88.0 (82.3)</td>
<td>23.0 (18.5)</td>
<td>6.8</td>
<td>6.9 (4 µs)</td>
<td>-7/9 (4 µs)</td>
</tr>
<tr>
<td>-10</td>
<td>264.3</td>
<td>23</td>
<td>4 ns to 8 s</td>
<td>10 s</td>
<td>71.5 (66.2)</td>
<td>27.4 (24.8)</td>
<td>5.0</td>
<td>4.8 (4 µs)</td>
<td>-9/11 (4 µs)</td>
</tr>
<tr>
<td>0</td>
<td>274.0</td>
<td>27</td>
<td>8 ns to 512 ms</td>
<td>10 s</td>
<td>75.1 (68.0)</td>
<td>27.1 (18.8)</td>
<td>5.4</td>
<td>5.7 (4 µs)</td>
<td>-9/13 (4 µs)</td>
</tr>
<tr>
<td>10</td>
<td>283.6</td>
<td>27</td>
<td>2 ns to 128 ms</td>
<td>4 s</td>
<td>82.2 (77.8)</td>
<td>30.3 (26.9)</td>
<td>5.2</td>
<td>3.9 (4 µs)</td>
<td>-10/11 (4 µs)</td>
</tr>
<tr>
<td>20</td>
<td>293.2</td>
<td>27</td>
<td>2 ns to 128 ms</td>
<td>4 s</td>
<td>81.7 (75.3)</td>
<td>20.1 (14.9)</td>
<td>6.8</td>
<td>6.9 (4 µs)</td>
<td>-7/7 (4 µs)</td>
</tr>
<tr>
<td>25</td>
<td>293.7</td>
<td>27</td>
<td>2 ns to 512 ms</td>
<td>2 s</td>
<td>78.3 (73.4)</td>
<td>25.5 (20.0)</td>
<td>5.3</td>
<td>5.2 (4 µs)</td>
<td>-11/12 (4 µs)</td>
</tr>
<tr>
<td>30</td>
<td>298.5</td>
<td>27</td>
<td>2 ns to 128 ms</td>
<td>2 s</td>
<td>83.5 (75.9)</td>
<td>25.3 (16.1)</td>
<td>5.6</td>
<td>6.8 (4 µs)</td>
<td>-11/10 (4 µs)</td>
</tr>
<tr>
<td>40</td>
<td>308.9</td>
<td>24</td>
<td>2 ns to 64 ms</td>
<td>2 s</td>
<td>84.7 (80.3)</td>
<td>30.6 (21.7)</td>
<td>4.7</td>
<td>7.1 (4 µs)</td>
<td>-9/10 (4 µs)</td>
</tr>
<tr>
<td>50</td>
<td>318.9</td>
<td>22</td>
<td>100 ns to 256 ms</td>
<td>4 s</td>
<td>77.0 (73.0)</td>
<td>31.3 (24.8)</td>
<td>4.6</td>
<td>4.8 (4 µs)</td>
<td>-9/9 (4 µs)</td>
</tr>
<tr>
<td>60</td>
<td>328.8</td>
<td>24</td>
<td>100 ns to 1 s</td>
<td>4 s</td>
<td>84.5 (79.3)</td>
<td>30.0 (18.1)</td>
<td>4.8</td>
<td>5.6 (2 µs)</td>
<td>-8/9 (2 µs)</td>
</tr>
<tr>
<td>70</td>
<td>338.8</td>
<td>21</td>
<td>100 ns to 1 s</td>
<td>4 s</td>
<td>84.6 (80.3)</td>
<td>30.6 (21.7)</td>
<td>4.7</td>
<td>4.8 (800 ns)</td>
<td>-6/6 (800 ns)</td>
</tr>
</tbody>
</table>

a Waiting time between two subsequent laser pulses  
b $R_{merge}$ = $\frac{\sum_{hkl} \sum_{j} I_{hkl,j} - \langle I_{hkl} \rangle}{\sqrt{\sum_{hkl} \sum_{j} I_{hkl,j}^{2} - \langle I_{hkl} \rangle^{2}}}$  
c $R_{scale}$ = $\frac{\sum_{hkl} F_{hkl}^{A} - F_{hkl}^{D}}{\sqrt{\sum_{hkl} F_{hkl}^{A}^{2} - F_{hkl}^{D}^{2}}}$

2.8.2 Time-resolved spectroscopy

2.8.2.1 Data collection: PYP in solution

A 400 nL drop of protein concentration 2-3 mg/mL is mounted as described in the Sec. 2.3.5. A time series of 10 time delays starting from 30 µs to 2 s at $T = 22$ °C is obtained.
The waiting time between two consecutive laser pulses is 5 s. The accumulations numbers and pulse widths are set for each time delay (Tab. 2.2).

### 2.8.2.2 Data collection: Crushed single PYP crystal

A crystal of size ~200 × 200 x 200 µm$^3$ is crushed between two cover slides as described in the Sec. 2.3.6. Data sets at two different temperatures 30$^\circ$C and 0$^\circ$C are collected. At 30$^\circ$C 24 time points covering the range from 25 µs to 2 s are obtained. Waiting time between two subsequent laser pulses is 5 sec. At 0$^\circ$C, a time range from 25 µs to 3s is covered by 34 time delays. Waiting time between two subsequent laser pulses is 10 s. For both temperatures, the accumulation numbers and pulse widths are different depending on the time delay (Tab. 2.2). The slower part of the photocycle is selected for all spectroscopic experiments as in the pH and dose experiments. Reaction initiation is achieved with 380 µJ of a 6 ns laser pulse at 446 nm.

**Table 2.2** Data collection parameters assigned to different time delays for time-resolved spectroscopic experiments.

<table>
<thead>
<tr>
<th>Time delay (range)</th>
<th>Pulse width</th>
<th>Accumulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>25µs - 80µs</td>
<td>20</td>
<td>45</td>
</tr>
<tr>
<td>81µs - 200µs</td>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>201µs - 1ms</td>
<td>60</td>
<td>20</td>
</tr>
<tr>
<td>1ms - 500ms</td>
<td>100</td>
<td>20</td>
</tr>
<tr>
<td>501ms - 3s</td>
<td>170</td>
<td>20</td>
</tr>
</tbody>
</table>

### 2.8.2.3 Data Analysis

Difference absorption spectra are obtained for solution (Fig. 3.12A) and crushed crystal (Fig. 3.13A-B) using the data analysis program described in Sec. 2.3.12. Spectral changes in the wavelength range from 400 to 520 nm are selected for the SVD analysis. rSVs obtained from these analysis are inspected for the presence of kinetic phases (Fig. 3.12B
& Fig. 3.13C-D). Relaxation times obtained from crushed crystal are compared at different temperatures to examine the effect of temperature on the photocycle.

References


3 Results

3.1 Interaction of Nitric Oxide (NO) with Catalase: X-ray crystallographic analysis

3.1.1 Crystal structures

All three crystals forms obtained are orthorhombic with space group P2\(_1\)2\(_1\)2\(_1\). The unit cell dimensions vary slightly depending on the crystal form, in the range of a = 83 ~ 86 Å, b = 139 ~ 140 Å, c = 228 ~ 229 Å and α = β = γ = 90° (Tab. 3.1). Multiple datasets are collected for each crystal form and the best one is picked based on high resolution, low degree of mosaicity, low R\(_{\text{sym}}\) values and high completeness. For these data sets, typical resolution extends to 1.9 Å and mosaicity is about ~ 0.3°. Data completeness is more than 90% with an R\(_{\text{sym}}\) < 10%. I/σ\(_I\) value found in the last resolution cell is 2.5.

Table 3.1 Data collection statistics for the three crystal forms of the catalase. Table from Purwar et al., 2011.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>CAT-NH(_3)</th>
<th>CAT-5</th>
<th>CAT-NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soaking</td>
<td>14 hours NH(_3)-free buffer</td>
<td>5 minutes 100 mM DEANO</td>
<td></td>
</tr>
<tr>
<td>a (Å)</td>
<td>83.48</td>
<td>83.22</td>
<td>86.11</td>
</tr>
<tr>
<td>b (Å)</td>
<td>140.72</td>
<td>140.92</td>
<td>139.94</td>
</tr>
<tr>
<td>c (Å)</td>
<td>229.52</td>
<td>229.37</td>
<td>228.02</td>
</tr>
<tr>
<td>α = β = γ (deg)</td>
<td>90</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Spacegroup</td>
<td>P2(_1)2(_1)2(_1)</td>
<td>P2(_1)2(_1)2(_1)</td>
<td>P2(_1)2(_1)2(_1)</td>
</tr>
<tr>
<td>Unit cell volume(Å(^3))</td>
<td>2696241</td>
<td>2689905</td>
<td>2747694</td>
</tr>
<tr>
<td>Resolution(Å)</td>
<td>1.99</td>
<td>1.90</td>
<td>1.88</td>
</tr>
<tr>
<td>R(_{\text{sym}}) (last shell)</td>
<td>9.4(30.3)</td>
<td>8.9(34.3)</td>
<td>9.1(32.9)</td>
</tr>
<tr>
<td>Completeness (last shell) (%)</td>
<td>92.8(92.8)</td>
<td>97.9(97.1)</td>
<td>92.2(89.9)</td>
</tr>
<tr>
<td>Redundancy (last shell)</td>
<td>5.7(5.5)</td>
<td>3.6 (3.3)</td>
<td>5.1 (4.3)</td>
</tr>
<tr>
<td>I/σ(_I) (last shell)</td>
<td>5.0(2.5)</td>
<td>5.1(2.5)</td>
<td>5.6(2.1)</td>
</tr>
<tr>
<td>R(<em>{\text{cry}})/R(</em>{\text{free}}) no water (%)</td>
<td>23.68/27.05</td>
<td>24.1/26.9</td>
<td>24.05/27.05</td>
</tr>
<tr>
<td>Water molecules</td>
<td>2126</td>
<td>1277</td>
<td>2013</td>
</tr>
<tr>
<td>R(<em>{\text{cry}})/R(</em>{\text{free}}) with water (%)</td>
<td>16.83/20.69</td>
<td>22.16/25.82</td>
<td>18.56/22.05</td>
</tr>
<tr>
<td></td>
<td>Heme A</td>
<td>Heme B</td>
<td>Heme C</td>
</tr>
<tr>
<td>------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>rmsd bond length</td>
<td>0.006</td>
<td>0.007</td>
<td>0.006</td>
</tr>
<tr>
<td>rmsd bond angle (deg)</td>
<td>1.3</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td>Ramachandran outliers</td>
<td>53 (2.9%)</td>
<td>62 (3.3%)</td>
<td>53 (2.9%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Occupation from electron count (%)</th>
<th>Heme A</th>
<th>Heme B</th>
<th>Heme C</th>
<th>Heme D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>70 6.0 51</td>
<td>71 3.7 40</td>
<td>82 1.4 35</td>
<td>95 1.7 74</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refined Occupancy (%)</td>
<td>Heme A</td>
<td>Heme B</td>
<td>Heme C</td>
<td>Heme D</td>
</tr>
<tr>
<td></td>
<td>82 55</td>
<td>100 52</td>
<td>77 45</td>
<td>101 61</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FE out of plane distance (Å)</td>
<td>Heme A</td>
<td>Heme B</td>
<td>Heme C</td>
<td>Heme D</td>
</tr>
<tr>
<td></td>
<td>0.063 0.14 0.016</td>
<td>-0.054 0.123 -0.035</td>
<td>-0.044 0.132 -0.020</td>
<td>-0.052 -0.005 0.019</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tilt angle (τ) / bend angle (φ) / tilt+bend angle (α) (deg)</td>
<td>Heme A</td>
<td>Heme B</td>
<td>Heme C</td>
<td>Heme D</td>
</tr>
<tr>
<td></td>
<td>7.98 9.14/19.6/28.74</td>
<td>6.01 8.17/5.03/3.13</td>
<td>0.51 12.28/11.78/24.06</td>
<td>5.33 8.23/9.72/17.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMS coordinate error (Å)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.181 0.262 0.214</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iron-Nitrogen distance (Å)</td>
<td>Heme A</td>
<td>Heme B</td>
<td>Heme C</td>
<td>Heme D</td>
</tr>
<tr>
<td></td>
<td>2.10 1.92</td>
<td>2.11 1.88</td>
<td>2.11 1.85</td>
<td>2.10 1.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nitrogen (NH3) –Water/ Oxygen (NO) -Water (Å)</td>
<td>Heme A</td>
<td>Heme B</td>
<td>Heme C</td>
<td>Heme D</td>
</tr>
<tr>
<td></td>
<td>3.05 2.42</td>
<td>3.05 2.69</td>
<td>2.96 2.58</td>
<td>2.76 2.52</td>
</tr>
<tr>
<td>Iron-TYR distance (Å)</td>
<td>3.0±0.13</td>
<td>2.55±0.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>----------</td>
<td>----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heme A</td>
<td>2.01</td>
<td>2.00</td>
<td>2.02</td>
<td></td>
</tr>
<tr>
<td>Heme B</td>
<td>2.00</td>
<td>1.92</td>
<td>1.99</td>
<td></td>
</tr>
<tr>
<td>Heme C</td>
<td>1.96</td>
<td>1.89</td>
<td>1.96</td>
<td></td>
</tr>
<tr>
<td>Heme D</td>
<td>2.04</td>
<td>2.01</td>
<td>1.96</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.0±0.03</td>
<td>2.0±0.06</td>
<td>1.98±0.025</td>
<td></td>
</tr>
<tr>
<td>iron – water distance(^b) (Å)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heme A</td>
<td>4.37</td>
<td>4.00</td>
<td>4.27</td>
<td></td>
</tr>
<tr>
<td>Heme B</td>
<td>4.41</td>
<td></td>
<td>4.20</td>
<td></td>
</tr>
<tr>
<td>Heme C</td>
<td>4.11</td>
<td></td>
<td>4.25</td>
<td></td>
</tr>
<tr>
<td>Heme D</td>
<td>4.10</td>
<td>3.98</td>
<td>4.21</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.25±0.17</td>
<td>3.99±0.01</td>
<td>4.23±0.029</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Root mean square deviation (RMSD) from ideal geometry, RMSDs of dihedrals and impropers were 23° and 1°, respectively, for all models. \(^b\)Distance from the heme iron to the closest water molecule near His 74

### 3.1.2 Ammonia-bound Catalase form (Cat-NH\(_3\))

A refined model with an \(R_{\text{cryst}}\) value of 23.7\% at 1.99Å is obtained. As shown in Fig. 1.2A, a catalase molecule has four subunits and each subunit contains a heme B as an active site. Previously-solved crystal structures, under the same conditions as ours, show that the BLC typically incorporates one tightly bound NADPH per subunit (Kirkman and Gaetani, 1984, Fita & Rossmann, 1985). However, our structure shows no electron density due to NADPH in the expected region (Fig. 1.2C). From this, we conclude that the NADPH is lost during purification by size exclusion. The total number of water molecules is around 2100 suggesting the catalase is a wet protein (Fig. 1.2D). A large number of water molecules are found in 140 cavities (Tab. 3.2). Adding all the waters to the structure, the R-factor is decreased to 16.8\%. For Cat-NH\(_3\), a strong positive electron density feature is present.
Figure 3.1 Structure of the heme pocket in the three different catalase forms. Water molecules within 8Å of the heme iron are displayed (red spheres). A sigmaA weighted mFobs-DFcalc simulated annealing difference map is shown in green (3.7σ). A. Cat-NH3 B. Cat-5 C. Cat-NO. All the relevant residues in the heme pocket are marked in panel B. Figure from Purwar et al., 2011.

near the heme iron (Fig. 3.1A). Since high concentrations of ammonium are present in the crystallization buffer, we propose that the electron density arises from an ammonia molecule bound to the heme iron. For each of the four hemes, the electron density at the 6th coordination site is integrated. The electron count at each active site is around ~ 9 e⁻, which supports the fact that the sites are fully occupied by NH₃. An NH₃ model is placed into the electron density at a distance of 1.9 Å using a bond length restraint of ~ 50 kcal mol⁻¹ Å⁻² and its position is refined. The occupancy of the NH₃ molecule refined to ~ 1.0 for all four subunits. The model precision is 0.18 Å as estimated by a Luzzati plot (Sec. 2.1.20). The Number of Ramachandran outliers is around ~ 53 (Tab. 3.1). For all four subunits, tilt angle of NH₃ molecule with respect to the heme plane is calculated (Tab. 3.1). An average value of the tilt angle is ~5°. For this crystal form, iron out of plane distance for each heme is approximately ± 0.05Å which indicates that the iron is almost in plane (Tab. 3.1).

3.1.3 The five-coordinate Cat-5
The final model of Cat-NH$_3$ is used as a starting model to refine Cat-5. After removing the ammonia molecule at the 6$^{th}$ coordination site of the iron, as well as the water molecules, the model is refined to 1.9Å (Tab. 3.1). Inclusion of water molecules lowered the R-value from 24.1% to 22.1%. About 1200 waters are found. Almost no positive density appeared at the heme active site (Fig. 3.1B) where NH$_3$ molecule is bound in the first crystal form Cat-NH$_3$. By integrating the faint electron density features around the sixth coordination site, the average electron count is less than 1 (Tab. 3.1). This demonstrates that NH$_3$ molecule is successfully removed from this crystal form, leaving a vacant 6$^{th}$ coordination site at the iron. The heme iron is 5-coordinate and therefore we call this crystal form Cat-5.

**Table 3.2 Number of cavities in some selected proteins.** Table from Purwar et al., 2011.

<table>
<thead>
<tr>
<th>Protein</th>
<th># of cavities</th>
<th>H$_2$O in cavity/total H$_2$O</th>
<th>%</th>
<th>PDB-entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catalase$^a$</td>
<td>140</td>
<td>100/1282</td>
<td>8</td>
<td>This work</td>
</tr>
<tr>
<td>Catalase (human, 1.5 Å)$^b$</td>
<td>108</td>
<td>197/2298</td>
<td>9</td>
<td>1DGF</td>
</tr>
<tr>
<td>Cytochrome-c oxidase$^c$</td>
<td>95</td>
<td>85/1579</td>
<td>5</td>
<td>3ABK</td>
</tr>
<tr>
<td>L29W-Mbdeoxy$^d$</td>
<td>2</td>
<td>1/100</td>
<td>1</td>
<td>2BL1</td>
</tr>
<tr>
<td>Mbdeoxy (to 1.15 Å)$^e$</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1A6N</td>
</tr>
</tbody>
</table>

$^a$ This study; $^b$ (Putnam et al., 2000); $^c$ (Ohta et al., 2010); $^d$ (Nienhaus et al., 2005), $^e$ (Vojtechovský et al., 1999)

### 3.1.4 Cat-NO Structure

The amount of NO needed to observe significant binding to the catalase crystals is optimized by varying the DEANO concentration from 10 to 200 mM. DEANO concentrations above 100 mM provide more than 50% NO occupancy in the heme pocket. Data are collected to 1.88Å (Tab. 3.1). The ammonia-free (Cat-5) structure without the water molecules is used as the initial model for the refinement. After inserting the waters, the R-value drops from 24.1% to 18.6%. For this crystal form, total
~2000 water molecules are found. Similar number of water molecules was observed for Cat-NH₃ form whereas for Cat-5, only 1200 waters were observed (Tab. 3.1). A positive electron density feature appears at the 6th coordination site of the heme iron (Fig. 3.1C), which was absent in Cat-5. This positive electron density is identified as a NO molecule. The refined NO occupancy is around ~ 55% on average (Tab. 3.1). A single water molecule is observed in the vicinity of the NO for all four hemes. This seems to be an important factor in defining the binding geometry of the NO with respect to the heme plane (see Discussion). The NO is slightly bent with respect to the heme plane as described above (Fig. 2.30B, Fig. 3.1C and Tab. 3.1). For the different subunits the bending angle varies between 5° and 20°, with an average of 12°. Note, that the bending angle strongly depends on the restraints employed. If the bending angle is allowed to vary in an unrestrained fashion, the bending angle increases up to ~ 70° (data not shown). The distance from the iron to the NO nitrogen refined to 1.85 Å, and the distance from the NO oxygen to the heme-pocket water is 2.5 Å. This indicates a strong hydrogen bond (Fig. 3.1C and Tab. 3.1).

Table 3.3 Tilt+bend angles for NO found in different heme-iron proteins with the iron in different oxidation states (Mb: myoglobin, NP4: nitrophorin 4). Table from Purwar et al., 2011.

<table>
<thead>
<tr>
<th>Nitrosyl complexes</th>
<th>tilt+bend [°]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mb(Fe²⁺)NO&lt;sup&gt;a&lt;/sup&gt;</td>
<td>60</td>
</tr>
<tr>
<td>Mb(Fe³⁺)NO&lt;sup&gt;a&lt;/sup&gt;</td>
<td>33</td>
</tr>
<tr>
<td>NP4(Fe²⁺)-NO&lt;sup&gt;b&lt;/sup&gt;</td>
<td>70/3</td>
</tr>
<tr>
<td>NP4(Fe³⁺)-NO&lt;sup&gt;c&lt;/sup&gt;</td>
<td>24</td>
</tr>
<tr>
<td>NP4(Fe⁴⁺)-NO&lt;sup&gt;d&lt;/sup&gt;</td>
<td>40</td>
</tr>
<tr>
<td>Catalase (Fe³⁺) (this study)</td>
<td>19</td>
</tr>
</tbody>
</table>

<sup>a</sup> (Copeland et al., 2003);  <sup>b</sup> (Weichsel et al., 2000);  <sup>c</sup> (Roberts et al., 2001),  <sup>d</sup> (Maes et al., 2005)
Table 3.4 Structural differences in Å relative to Cat-5 averaged over all 4 subunits. Table from Purwar et al., 2011.

<table>
<thead>
<tr>
<th></th>
<th>Cat-NH₃</th>
<th>Cat-NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fe out of plane and heme nitrogens [Å]</td>
<td>0.1±0.01</td>
<td>0.1±0.01</td>
</tr>
<tr>
<td>proximal side of the heme [Å]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TYR 357-O₃</td>
<td>0.11±0.02</td>
<td>0.14±0.04</td>
</tr>
<tr>
<td>distal side of the heme [Å]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIS 74-N₁/N₂</td>
<td>0.12±0.06</td>
<td>0.14±0.06</td>
</tr>
<tr>
<td>ASN 147-O₂/N₁/N₁₂</td>
<td>0.12±0.07</td>
<td>0.16±0.08</td>
</tr>
<tr>
<td>PHE 160-Cζ</td>
<td>0.15±0.03</td>
<td>0.13±0.07</td>
</tr>
<tr>
<td>Total RMS differences⁷</td>
<td>0.34±0.04</td>
<td>0.37±0.05</td>
</tr>
</tbody>
</table>

⁷ The structures were superimposed by a least-squares fit (Xfit) (McRee, 1999).

3.2 Interaction of Nitric Oxide (NO) with Catalase: Microspectrophotometric analysis

3.2.1 Spectral changes for Cat-5, Cat-NH₃ and Cat-NO

For Cat-5 a total of 60 absorption spectra $A(\lambda)_{Δt}$ and the corresponding $ΔA(\lambda)_{Δt}$ and $ΔA_{int}(t)$ are plotted in Fig. 3.2A, 3.2D & 3.2G, respectively. Spectral changes that are characteristic to the reduction of heme iron from Fe(III) (met) to Fe(II) (deoxy) appear for exposure time over ~100s (Fig. 3.2A & 3.2D)(Shimizu et. al., 1988). On longer timescales the area of the integrated difference spectrum $ΔA_{int}(t)$ decreases which may indicate crystal degradation (Fig. 3.2G). When the experiment is repeated with other crystals of the Cat-5 form, the rate of initial rise remains fairly invariant. However, the rate of subsequent decay varies significantly from one experiment to the other (data not shown). For Cat- NH₃, $A(t_{i})$, $ΔA(t_{i})$ and $ΔA_{int}(t_{i})$ are plotted in panel B, E and H, respectively. The observed spectral changes for Cat-NH₃ are bit more complicated than for Cat-5, but they still exhibit features characteristic of heme reduction (Fig. 3.2B & 3.2E). Panels C, F and I represent $A(t_{i})$, $ΔA(t_{i})$ and $ΔA_{int}(t_{i})$ for Cat-NO, respectively.
Interestingly the broadening of the signal is observed in this crystal form (Fig. 3.2 C & F). A similar type of behavior is observed when nitrosylated myoglobin (Hoshino et. al., 2011).
al., 1996) or nitrophorin (Weichsel et al., 2005) are reduced. Note that almost no decay is observed in $\Delta A(t)$ (Fig. 3.2I) whereas this decay is immediately evident for Cat-5 as well as for Cat-NH$_3$ (Fig. 3.2H & I). Our spectrophotometric data show that catalase is reduced after 10 minutes while illuminated with X-ray beam.

### 3.3 Time-resolved Laue crystallography: Effect of X-ray dose on PYP

#### 3.3.1 Data processing

Laue data are processed up to 1.6 Å (Tab. 3.5). Since the data quality is poor in the last resolution shell, the difference maps are calculated to 1.8 Å. In order to address the radiation damage, the mean intensity $\langle I \rangle$, and the mean of the ratio of the intensity over its experimental uncertainty, $\langle I/\sigma_I \rangle$, in the resolution shell 1.9 to 1.8 Å are calculated. We used unscaled raw reflection intensities derived from the integration of the Laue spots.

#### 3.3.2 Average absorbed dose

For the PYP1 crystal, the absorbed dose by the X-ray illuminated voxel ($d_{\text{vox}}$) is $0.244 \times 10^4$ Gy (Schmidt et al., 2012). The dose deposited to the total irradiated crystal volume is $0.0414 \times 10^4$ Gy (Eq. 2.28). Average X-ray dose per time series is $3.3 \times 10^5$ Gy. The total absorbed dose ($D_{\text{nc}}$) in all 12 time series is ~ 4 MGy. In Fig. 3.3A, the $\langle I \rangle$ and $\langle I/\sigma_I \rangle$ values (solid and open squares, respectively) from all the datasets collected at time delay 32 ms are plotted as a function of uncorrected absorbed dose. The $\langle I \rangle$ and $\langle I/\sigma_I \rangle$ are fitted by a single exponential and, the half value for uncorrected dose ($D_{1/2, \text{nc}}$) is measured (Fig. 3.3A). $D_{1/2, \text{nc}}$ is more than $18.4 \times 10^5$ Gy when the $\langle I \rangle$ values are used and is ~ 20 x $10^5$ Gy when the $\langle I/\sigma_I \rangle$ values are used (Fig. 3.3A & Tab. 3.6).
### Table 3.5 Data statistics for 1\textsuperscript{st}, 6\textsuperscript{th} and 12\textsuperscript{th} time-series after analysis with Precongnition/Epinorm. Table from Schmidt et al., 2012.

<table>
<thead>
<tr>
<th>Short time-series</th>
<th>1\textsuperscript{st}</th>
<th>6\textsuperscript{th}</th>
<th>12\textsuperscript{th}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>dark 2ms 32ms</td>
<td>dark 2ms 32ms</td>
<td>dark 2ms 32ms</td>
</tr>
<tr>
<td>time-delay</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completeness (%)(^a)</td>
<td>80.05 73.58 79.34</td>
<td>76.38 73.8 75.72</td>
<td>68.57 67.84 67.86</td>
</tr>
<tr>
<td>Completeness (%)(^b)</td>
<td>66.88 55.97 66.39</td>
<td>59.81 54.62 58.18</td>
<td>44.96 43.40 43.93</td>
</tr>
<tr>
<td>R(_{\text{merge}}) on F(^c) (%)</td>
<td>5.74 6.01 5.75</td>
<td>6.02 6.83 6.30</td>
<td>7.93 7.83 8.14</td>
</tr>
<tr>
<td>24.83 24.05 25.21</td>
<td>22.9 20.92 22.32</td>
<td>18.59 18.74 17.52</td>
<td></td>
</tr>
<tr>
<td>R(_{\text{scale}}) 1.6-1.64</td>
<td>11.7 11.1 13.2 12.7</td>
<td>15.3 14.9</td>
<td></td>
</tr>
<tr>
<td>(^a) Resolution range: 1.00–1.6Å, (^b) resolution range: 1.87–1.6Å.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.3.3 Correction to the absorbed dose (\(\text{D}_{1/2, \text{corr}}\))
For the actual data, the total shared volume \( V_c \) owing to subsequent crystal settings is 77\% (Tab. 3.6) and therefore 23\% of fresh volume on an average is exposed in different crystal settings (Sec. 2.2.6, Fig. 2.31B) (Schmidt et al., 2012). For the first 5 time series, vertical displacements \( \langle y' \rangle \) are negligible (Tab. 3.6). Large values of \( \langle y' \rangle \) are observed from 5\textsuperscript{th} to 12\textsuperscript{th} time series and a total 25\% of the X-ray probed volume is freshly exposed. The corrected X-ray doses \( D_{corr} \) are obtained (Fig. 3.3B & Tab. 3.6).

### 3.3.4 Correction for laser damage using the control experiment

The thickness of the crystal used for laser experiment is 150\,\mu m. The common shared volume \( V_c \) is 80\% which is slightly larger than that obtained for real time series (77\%). No correction is needed due to vertical displacements \( \langle y' \rangle \) of the X-ray beam because crystal positions remain fairly stable in all dark X-ray exposures (data not shown). The corrected virtual doses \( D_L \) are obtained (Fig. 3.3B). The \( \langle I \rangle^L \) (solid black triangles) values are plotted as a function of \( D_L \) and fitted by a straight line with the slope \( S_L = -0.0113 \times 10^{-5} \langle I \rangle/\text{Gy} \) (Fig. 3.3B). The virtual half dose value \( D_{1/2}^I \) (not shown), at which the normalized intensity reaches half of its maximum value, is about 4.5 Mg. The intensity free from the laser damage, \( \langle I \rangle_{\text{free}} \), is calculated (Eq. 2.30) and plotted as a function of corrected absorbed dose \( D_{corr} \) (Fig. 3.3B). The \( \langle I \rangle_{\text{free}} \) values (red solid squares) as are fitted by a straight line as well as by a single exponential. The crystallographic dose limit is given by \( D_{1/2}^{corr} = 16.8 \times 10^5 \text{ Gy} \), which is same whether \( \langle I \rangle_{\text{free}} \) or \( \langle I/\sigma_I \rangle_{\text{free}} \) are used (Fig. 3.3B & Tab. 3.7).
Table 3.6 Absorbed doses (in $10^5$ Gy) for 12 consecutive time-series, before and after corrections. N is the number of reflections in the resolution shell 1.9-1.8 Å. ($I$)$_{32ms}$ and ($I/\sigma_I$)$_{32ms}$ are determined from the 32 ms time point in the resolution shell 1.9-1.8 Å. ($I$)$_{free}$ and ($I/\sigma_I$)$_{free}$ are determined from dose data after applying laser corrections; Table from Schmidt et al., 2012.

<table>
<thead>
<tr>
<th>time-series</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average dose N$^a$</td>
<td>3.3</td>
<td>6.6</td>
<td>9.9</td>
<td>13.2</td>
<td>16.5</td>
<td>19.8</td>
<td>23.1</td>
<td>26.4</td>
<td>29.7</td>
<td>33.0</td>
<td>36.3</td>
<td>39.6</td>
</tr>
<tr>
<td>Dcorr$^1$ (V$_c$=77%)</td>
<td>2.54</td>
<td>5.08</td>
<td>7.62</td>
<td>10.16</td>
<td>12.71</td>
<td>15.25</td>
<td>17.79</td>
<td>20.32</td>
<td>22.87</td>
<td>25.41</td>
<td>27.95</td>
<td>30.49</td>
</tr>
<tr>
<td>(\langle y \rangle_{[\mu m]})</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>8</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>Dcorr$_2$ (V$_c$ + (\langle y \rangle))</td>
<td>2.54</td>
<td>5.08</td>
<td>7.62</td>
<td>10.16</td>
<td>12.29</td>
<td>13.98</td>
<td>15.41</td>
<td>16.59</td>
<td>18.30</td>
<td>19.90</td>
<td>21.42</td>
<td>22.36</td>
</tr>
<tr>
<td>N$_{reflections}$</td>
<td>3852</td>
<td>3860</td>
<td>3861</td>
<td>3861</td>
<td>3859</td>
<td>3853</td>
<td>3864</td>
<td>3889</td>
<td>3908</td>
<td>3888</td>
<td>3894</td>
<td></td>
</tr>
<tr>
<td>($I$)$_{32ms}$</td>
<td>330</td>
<td>297</td>
<td>271</td>
<td>237</td>
<td>211</td>
<td>185</td>
<td>160</td>
<td>151</td>
<td>123</td>
<td>110</td>
<td>88</td>
<td>84</td>
</tr>
<tr>
<td>($I/\sigma_I$)$_{32ms}$</td>
<td>11.07</td>
<td>10.22</td>
<td>9.35</td>
<td>8.23</td>
<td>7.33</td>
<td>6.50</td>
<td>5.71</td>
<td>5.31</td>
<td>4.42</td>
<td>3.98</td>
<td>3.33</td>
<td>3.26</td>
</tr>
<tr>
<td>($I/\sigma_I$)$_{free}$</td>
<td>11.31</td>
<td>10.67</td>
<td>9.97</td>
<td>9.00</td>
<td>8.10</td>
<td>7.28</td>
<td>6.47</td>
<td>6.07</td>
<td>5.12</td>
<td>4.66</td>
<td>3.94</td>
<td>3.89</td>
</tr>
<tr>
<td>(\tau_{SV1}[\text{ms}])</td>
<td>22</td>
<td>20</td>
<td>26</td>
<td>30</td>
<td>24</td>
<td>21</td>
<td>29</td>
<td>35</td>
<td>27</td>
<td>25</td>
<td>18</td>
<td>24</td>
</tr>
<tr>
<td>(\tau_{peak}[\text{ms}] / (1/k_2))</td>
<td>23</td>
<td>23</td>
<td>27</td>
<td>30</td>
<td>48</td>
<td>65</td>
<td>95</td>
<td>89</td>
<td>168</td>
<td>136</td>
<td>113</td>
<td>nd</td>
</tr>
<tr>
<td>R$^{ad}$</td>
<td>6.04</td>
<td>7.27</td>
<td>5.72</td>
<td>6.12</td>
<td>3.91</td>
<td>3.04</td>
<td>2.49</td>
<td>2.45</td>
<td>2.07</td>
<td>1.8</td>
<td>1.57</td>
<td>1.53</td>
</tr>
</tbody>
</table>

$^a$not corrected, $^1$dose correction (D$_{corr}$) due to crystal settings $^2$correction due vertical displacements are added to the D$_{corr}$, nd not detected

3.3.5 First Right singular vectors (rSVs) from SVD

After the SVD analysis, only one significant right singular vector (rSV1) for all 12 time series is found. Here, the rSVs for the 1$^{st}$ and 10$^{th}$ time-series are shown (Fig. 3.4). The rSVs for the first time course (rSV1(1$^{st}$), black solid spheres) is fitted by the sum of two exponentials, one for a rising phase and another for a decaying phase with a relaxation time $\tau(1)_{SV1}$ (Fig. 3.4A). For comparison, rSV1(1$^{st}$) is fitted with one exponential with the same characteristic time $\tau(1)_{SV1}$. In the 10$^{th}$ time-series, a rising phase is barely observed in rSV1(10$^{th}$) (Fig. 3.4B).
Consequently, only one exponential with \( \tau (1) \) is used for the fit. Both \( \tau (1) \) and \( \tau (10) \) are very similar, however the amplitudes and the offsets of \( rSV1(1) \) and \( rSV1(10) \) differ grossly (compare Fig. 3.4A & B).

### 3.3.6 The ratio of the amplitude to the offset

The level of noise in the difference electron density maps becomes higher with increasing radiation damage. As observed, the offset in \( rSV1(10) \) is larger than \( rSV1(1) \) (Fig. 3.4). This effect has also been noticed in spectroscopy (Henry & Hofrichter, 1992).
From the first rSV (rSV1), we calculate the ratio of the amplitude to the offset (Fig. 3.5A, Tab. 3.6). This ratio is then computed for the rSV1s of all the 12 time-series. The calculated values are plotted as a function of dose and fitted by a single exponential (black spheres) (Fig. 3.5A). From the fit, the half-value, \( D_{1/2}^K \), is determined and the corresponding dose is the kinetic half-value, \( D_{1/2}^K = 9.7 \times 10^5 \) (Fig. 3.5A & Tab. 3.7). The rSV1s for all the time-courses are plotted in the three dimensions (Fig. 3.5B). The orange regime ends after 72 data sets. The red line indicates that posterior analysis of the data beyond this dose will not be possible. Figure from Schmidt et al., 2012.

**Figure 3.5 A.** Normalized \( R^{AO} \) values (black spheres) for rSV1 of all 12 time series as a function of adjusted dose. Black dashed curve: data fitted by a single exponential; Horizontal and vertical dashed line: kinetic dose limit  

**B.** Three dimensional plot of rSV1 for all short time-series as a function of corrected dose. The offsets in the rSV1s are indicated by the colored dotted lines. Green dotted line: small offsets; orange dotted line: offset increases slightly; red dotted line: offset increases strongly. The orange regime ends after 72 data sets. The red line indicates that posterior analysis of the data beyond this dose will not be possible. Figure from Schmidt et al., 2012.

### 3.3.7 Relaxation times
Table 3.7 and derived from uncorrected and corrected dose data. Table from Schmidt et al., 2012.

<table>
<thead>
<tr>
<th></th>
<th>(Gy)</th>
<th>(Gy, Sec. 3.3.6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(&lt;I&gt;); uncorrected dose</td>
<td>18.4 x 10^5</td>
<td></td>
</tr>
<tr>
<td>(&lt;I&gt;/\sigma_I&gt;); uncorrected dose</td>
<td>19.7 x 10^5</td>
<td></td>
</tr>
<tr>
<td>(&lt;I&gt;_{\text{free}}); corrected dose</td>
<td>16.8 x 10^5</td>
<td></td>
</tr>
<tr>
<td>(&lt;I&gt;/\sigma_{I_{\text{free}}}); corrected dose</td>
<td>16.8 x 10^5</td>
<td></td>
</tr>
<tr>
<td>(R^{I};) uncorrected dose (not shown in Fig. 3.5A)</td>
<td></td>
<td>16.0 x 10^5</td>
</tr>
<tr>
<td>(R^{I_{\text{free}}};) corrected dose</td>
<td></td>
<td>9.6 x 10^5</td>
</tr>
</tbody>
</table>

The relaxation times \((\text{red solid squares})\) extracted from the first rSVs and those calculated from the posterior analysis \((\text{green solid squares})\) are plotted (Fig. 3.6A).

Up to the \(= 9.7 \times 10^5\) Gy, the relaxation times agree with relaxation times .

After , and start to diverge as shown in Fig. 3.6A & B. For higher absorbed dose such as \(~22 \times 10^5\) Gy, the time scale of the reaction from the SVD analysis...
is still 26 ms, however the time scale from the posterior analysis reaches up to 140 ms (Fig. 3.6).

### 3.4 Time-resolved Laue crystallography: Effect of pH on PYP

#### 3.4.1 Data processing

Data are processed up to 1.6 Å. The completeness of the data to this resolution is acceptable and excellent up to 1.8 Å. Difference maps are calculated to 1.8 Å. Data collection statistics at pH 4, 7 and 9 for some selected time points are shown (Tab. 3.8). $R_{\text{merge}}$ on $|F|^2$ is around 4.5% for pH 7, < 6% for pH 4 and ~5% for pH 9, showing the excellent data quality. $R_{\text{scale}}$ is obtained when time-dependent data sets are scaled to the dark (reference) data at a resolution of 1.6 Å. For all three pH’s $R_{\text{scale}}$ are compared (Tab. 3.8).

**Table 3.8** Data collection statistics for PYP at three different pH conditions. $\Delta \rho$ denote highest positive and lowest negative difference electron density features in the unit of $\sigma$ value that is used for the difference maps. Table from Tripathi et al., 2012.

<table>
<thead>
<tr>
<th></th>
<th>pH 7</th>
<th>pH 4</th>
<th>pH 9</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>time delay ($t_i$)</strong></td>
<td>4µs</td>
<td>512µs</td>
<td>512ms</td>
</tr>
<tr>
<td>Completeness (%) $^a$</td>
<td>70.5</td>
<td>70.4</td>
<td>71.5</td>
</tr>
<tr>
<td>Completeness (%) $^b$</td>
<td>38.3</td>
<td>37.5</td>
<td>40.8</td>
</tr>
<tr>
<td>$\langle 1/\sigma_i \rangle$</td>
<td>29.8</td>
<td>26.5</td>
<td>29.5</td>
</tr>
<tr>
<td>$R_{\text{merge}}$ on $</td>
<td>F</td>
<td>^2$ (%)</td>
<td>4.6</td>
</tr>
<tr>
<td>$R_{\text{scale}}$ (%) $^b$</td>
<td>11.2</td>
<td>12.8</td>
<td>10.1</td>
</tr>
<tr>
<td>$\langle \Delta F \rangle$ (e$^-$)</td>
<td>9.1</td>
<td>9.9</td>
<td>6.7</td>
</tr>
<tr>
<td>$\langle \sigma_{\Delta F} \rangle$</td>
<td>4.2</td>
<td>4.8</td>
<td>4.1</td>
</tr>
<tr>
<td>$\Delta \rho_{\text{min}}/\sigma_{\Delta \rho}$</td>
<td>-11.2</td>
<td>-6.8</td>
<td>-4.7</td>
</tr>
<tr>
<td>$\Delta \rho_{\text{max}}/\sigma_{\Delta \rho}$</td>
<td>12.1</td>
<td>6.6</td>
<td>4.9</td>
</tr>
</tbody>
</table>

$^a$Resolution range 100 – 1.6 Å  $^b$in the last resolution shell 1.6-1.64 Å
Figure 3.7 A-C. Right singular vectors obtained from the SVD analysis at three different pH’s. 1st rSV: black spheres, 2nd rSV: red squares, 3rd rSV: green triangles. Two significant rSVs are present and each of them is fitted with sum of two exponentials. Fits for the 1st rSV (black solid curve) and the 2nd rSV (red solid curve) are shown. Relaxation times of the pR (τ1) and pB (τ2) determined from the fits are marked by vertical black dashed lines. A. at pH 7, 1st and 2nd rSVs are fitted by a sum of three exponentials. The third phase indicates a third intermediate with a relaxation time of 71 ms (gray dashed vertical line); B. at pH 4, C. at pH 9. D-F. Concentration profiles obtained from the posterior analysis using the proposed chemical kinetic mechanism (Fig. 2.17B). Relative concentration of intermediate pR (red), pB (blue), and pG (black) at D. pH 7.0, E. pH 4.0 and F. pH 9.0. Relaxation times τP (vertical dashed lines) calculated from microscopic rate coefficients k, are shown. Figure from Tripathi et al., 2012.
The rSVs obtained from the SVD analysis of the difference maps are plotted as a function of time (Fig. 3.7). Two significant rSVs, rSV1 (black solid spheres) and rSV2 (red solid squares), are present showing that two kinetic processes occur at each of three pH values (Fig. 3.7). Note, at pH 7, a small 3rd phase might be present (Fig. 3.7B), and at pH 9, the 2nd process is almost absent (Fig. 3.7C). The rSVs are fitted by sum of two exponentials except at pH 7 for which fitting is also tried with 3 exponentials. The relaxation times, $\tau_1$ and $\tau_2$ obtained from the fit correspond to pR and pB, respectively (Ihee et al., 2005).

### 3.4.2 Averaged difference maps and extrapolated electron density maps of intermediates

For both pH 4 and 7, two kinetic processes corresponding to the pR and pB are present (Fig. 3.7A-B & 3.7D-E). The average difference maps for pR and pB are obtained and corresponding $\Delta F_{\text{lin}}$ are calculated as described in Sec. 2.2.14. $N_c$s are determined for both pH as well as for both intermediates (Sec. 2.2.14). For example, at pH 7 the values of $N_c$ are about 15 and 25 for the pR and the pB states respectively. Using the $N_c$s, extrapolated maps are obtained (Fig. 3.8I.A & 3.8I.E). These maps are interpreted by atomic models which are subsequently refined and improved phases $\phi_{\text{new}}$ are obtained. Using $\phi_{\text{new}}$ and the measured difference structure factor amplitudes, new difference maps (Fig. 3.8D & 3.8H) as well as new extrapolated maps (Fig. 3.8I.C & I.G) are calculated. At pH 9 only one kinetic process corresponding to intermediate pR is clearly identified with $\tau_1$ of 1.5 ms (Fig. 3.7C & 3.7F). An extrapolated map for the pR is calculated as described for the other pHs (Tripathi et al., 2012). At this pH the 2nd process corresponding to pB is barely visible. The signal to noise ratio is too low in the difference map to obtain a meaningful extrapolated map. A very weak signal is found at the
Figure 3.8 Extrapolated and difference electron density maps at different pHs. I. pH 7; II. pH 4 and III. pH 9. pR (left panels, A-D) and pB (right panels, E-H); Panel A, C, E & G denotes extrapolated electron density maps at 1σ level (blue contour lines). Atomic models of intermediate states (red lines) are shown. Panels B, D, F & G show difference electron density maps; positive (blue contour lines) and negative (red contour lines) difference electron density are contoured at ±4σ. The dark state (pG) model (in yellow) is shown. Extrapolated maps before (panel A & E) and after (panel C & G) phase improvement. Difference electron density maps before (panel B & F) and after (panel D & H) phase improvement. The (difference) electron density features α and β are indicated by the arrow in E and G. Figure from Tripathi et al., 2012.

chromophore foot and also at Arg-52 that may suggest the presence of pB with very low
3.4.3 Structural analysis of intermediates at pH 7, pH 4 and pH 9

Data collection and refinement statistics for intermediates, pR and pB are quite similar for all pH values (Tab. 3.8). In our case, the pR intermediate adopts two conformations (i) pR<sub>CW</sub> and (ii) pR<sub>E46Q</sub> as reported previously (Ihee <i>et al.</i>, 2005). For all three pH’s, the relative occupancy of pR<sub>E46Q</sub> is lower in compared to the pR<sub>CW</sub>. In pR<sub>CW</sub>, the hydrogen bonds (H-bonds) of the chromophore phenolic oxygen with Tyr-42 and Glu-46 are preserved. The bond length for Tyr-42 is 2.9 Å and for Glu-46 is 2.7 Å at pH 7 whereas these lengths are the same for both residues at pH 4 (2.8 Å) and pH 9 (2.9 Å). In pR<sub>E46Q</sub>, the H-bond with Glu-46 is broken while that to Tyr-42 is preserved at a distance of 2.8 Å for all three pHs. In the case of pB, the H-bonds to Tyr-42 and Glu-46, both are broken. The chromophore head and Arg-52 are exposed to the protein solvent. In the extrapolated and difference density maps of the pB state, a feature (α) is detected close to chromophore phenolic oxygen, both at pH 7 and 4 (Fig. 3.8I.E & I.G; 3.8II.E & II.G). The feature α is interpreted as a water molecule which is H-bonded to the phenolate oxygen of the chromophore. The corresponding H-bond distance is 2.6 Å for pH 7 and 2.8 Å for pH 4. At pH 4, for state pB another large feature β appears in the chromophore pocket near Glu-46 and Tyr-42 (Fig. 3.8I.E & I.G; 3.8II.E & II.G). This feature has never been observed in any time-resolved crystallographic difference map on PYP. It is also interpreted as an additional water molecule which is H-bonded to Tyr-O<sub>η</sub> with a distance of 2.7 Å and weakly boned to Glu-46 O<sub>ε1</sub> with 3.3 Å.
3.4.4 Kinetic Analysis

The relaxation times ($\tau_1$) obtained from fitting of the rSVs are shown in Tab. 3.9. No significant differences are observed in $\tau_1$ for the pR state at pH 4 and pH 7. However, $\tau_1=1.5$ ms is significantly slower at pH 9. The relaxation time ($\tau_2$) for pB to dark state transition is more than 52 fold longer at pH 4 as compared to pH 7 (Tab. 3.9). The rate coefficients $k_1$, $k_2$ and $k_3$ determined from the posterior analysis are shown in Tab 3.9. At pH 7, the rate coefficient $k_3$ is much smaller than $k_1$. This suggests that the pathway from pR directly to pG is irrelevant at pH 7. At pH 4, $k_3$ is more than 20% of $k_1$ which indicates that a fraction of the molecules relaxes directly from pR to pG. At pH 9, the low occupation of pB did not allow the refinement of $k_1$ and $k_2$. The value of $k_3 = 494$ s$^{-1}$ implies that pR mainly decays directly to the pG, and pB is almost absent.

**Table 3.9** Relaxation times from SVD analysis ($\tau_1$ & $\tau_2$). Rate coefficients ($k_1$, $k_2$ & $k_3$) and their inverse, $\tau_p$ from the posterior analysis. Table from Tripathi et. al., 2012.

<table>
<thead>
<tr>
<th></th>
<th>pH=7</th>
<th>pH=4</th>
<th>pH=9*</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_1$</td>
<td>175 (+/- 8) µs</td>
<td>208 (+/- 12) µs</td>
<td>1.5 (+/-0.1) ms</td>
</tr>
<tr>
<td>$\tau_2$</td>
<td>9.2 (+/-0.2) ms</td>
<td>478 (+/-0.2) ms</td>
<td>20 (+/-0.1) ms</td>
</tr>
<tr>
<td>$k_1$</td>
<td>9730 (+/- 9.58) s$^{-1}$</td>
<td>13800 (+/- 3.12) s$^{-1}$</td>
<td>0.0</td>
</tr>
<tr>
<td>$k_2$</td>
<td>55.8 (+/- 1.02) s$^{-1}$</td>
<td>0.475 (+/- 0.002) s$^{-1}$</td>
<td>0.0</td>
</tr>
<tr>
<td>$k_3$</td>
<td>0.0</td>
<td>2940 (+/- 5.62) s$^{-1}$</td>
<td>494 (+/- 51.5) s$^{-1}$</td>
</tr>
<tr>
<td>$\tau_p$</td>
<td>$\tau_p = 1/k_2 \approx 18 ms$</td>
<td>$\tau_p = 1/(k_1 + k_3) \approx 60 ms$</td>
<td>$\tau_p = 1/k_3 \approx 2 ms$</td>
</tr>
</tbody>
</table>

* At this level of pB occupation, refinement of $k_1$ and $k_2$ is not possible.

3.5 Effect of temperature on PYP photocycle

3.5.1 Time-resolved Laue crystallography

3.5.1.1 Kinetic analysis

Data collection statistics are shown for all 14 temperature settings in Tab. 2.1.
Table 3.10 Relaxation times obtained ($\tau_i$) from the fit of rSVs for all temperature settings. The macroscopic coefficients ($\Lambda_i$) are the inverse of $\tau_i$. The temperature dependence of $\Lambda_i$ are fitted by the Van’t Hoff-Arrhenius equation. From these fitting, energy of activations are calculated.

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>$\tau_1$</th>
<th>$\tau_2$</th>
<th>$\tau_3$</th>
<th>$\tau_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-40</td>
<td>11 ns</td>
<td>1.9 µs</td>
<td>5.5 ms</td>
<td>1.4 s</td>
</tr>
<tr>
<td>-30</td>
<td>2 ns</td>
<td>120 ns</td>
<td>7 ms</td>
<td>420 ms</td>
</tr>
<tr>
<td>-20</td>
<td>2 ns</td>
<td>100 ns</td>
<td>5 ms</td>
<td>411 ms</td>
</tr>
<tr>
<td>-15</td>
<td>n.o.</td>
<td>36 ns</td>
<td>2.6 ms</td>
<td>537 ms</td>
</tr>
<tr>
<td>-10</td>
<td>n.o.</td>
<td>40 ns</td>
<td>3 ms</td>
<td>129 ms</td>
</tr>
<tr>
<td>0</td>
<td>n.o.</td>
<td>58 ns</td>
<td>697 µs</td>
<td>73 ms</td>
</tr>
<tr>
<td>10</td>
<td>n.o.</td>
<td>31 ns</td>
<td>300 µs</td>
<td>28 ms</td>
</tr>
<tr>
<td>20</td>
<td>n.o.</td>
<td>24 ns</td>
<td>198 µs</td>
<td>9 ms</td>
</tr>
<tr>
<td>25</td>
<td>n.o.</td>
<td>15 ns</td>
<td>198 µs</td>
<td>12 ms</td>
</tr>
<tr>
<td>30</td>
<td>n.o.</td>
<td>10 ns</td>
<td>99 µs</td>
<td>6 ms</td>
</tr>
<tr>
<td>40</td>
<td>n.o.</td>
<td>7 ns</td>
<td>41 µs</td>
<td>4.4 ms</td>
</tr>
<tr>
<td>50</td>
<td>n.o.</td>
<td>n.o.</td>
<td>40 µs</td>
<td>1.3 ms</td>
</tr>
</tbody>
</table>

Macroscopic rate coefficients

<table>
<thead>
<tr>
<th>$\Lambda_1$</th>
<th>$\Lambda_2$</th>
<th>$\Lambda_3$</th>
<th>$\Lambda_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>n.a.</td>
<td>1.6 x 10^{14}</td>
<td>1.9 x 10^{13}</td>
<td>7.7 x 10^{10}</td>
</tr>
</tbody>
</table>

Pre-factor $v$ [1/s]

| Energy of activation $E_a$ [kJ/mol] | 35.9 | 53.4 | 49.6 |

n.o. not observed, n.a. not applicable

Figure 3.9 Chemical kinetic mechanism proposed for the PYP photocycle. Intermediates, $I_T$, $I_{CT}$, $pR_1$, $pR_2$, $pB_1$ and $pB_2$ are shown in different colors. Associated rate coefficients $k_i$ are also shown. Figure from Schmidt et al., 2013.

The rSVs extracted from the SVD analysis for all temperatures are plotted as a function of employed time delays (Appendix E). The nomenclature for the PYP photocycle intermediates such as $I_T$, $I_{CT}$, $pR_1$, $pR_2$, $pB_1$ and $pB_2$ (Fig. 3.9) is same as described in...
previous works (Genick et al., 1997a, Ihee et al., 2005, Kim et al., 2012, Jung et al., 2013, Yeremenko et al., 2006). At -40 °C, three significant rSVs are found which are globally fitted with 4 exponentials as shown in Fig. 3.10A. Correspondingly, four kinetic processes with relaxation times $\tau_1 = 11$ ns, $\tau_2 = 1.9$ µs, $\tau_3 = 5.5$ ms, $\tau_4 = 1.4$ s are observed. The process with $\tau_1 = 11$ ns results from the non-zero laser pulse width and the decay of $I_T$ to both $I_{CT}$ and $pR_1$ that can be identified in the earliest difference maps. $\tau_2 = 1.9$ µs results from the relaxation of state $I_{CT}$ to $pR_2$, $\tau_3 = 5.5$ ms from the joint relaxations of $pR_1$ and $pR_2$ to $pB$ and $pG$ and $\tau_4 = 1.4$ s from $pB$ to $pG$ (Fig. 3.10A). These processes accelerate at higher temperatures. For comparison, relaxation times at -40 °C (black colored numbers) and at 25 °C (red colored numbers in bracket) are shown in Fig. 3.10A.
The relaxation rates $\Lambda_i = 1/\tau_i$ are plotted as a function of different temperature settings employed in the experiment (Fig. 3.10B-D). The rates are fitted by the Van't Hoff-Arrhenius equation (Eq. 1.4). From the fitting, the pre-factors and the energies of activation are determined (Tab. 3.10).

3.5.2 Time-resolved spectroscopy

3.5.2.1 Dark absorption spectra of PYP in solution and on crystalline slurry

![Absorption spectra](image)

**Figure 3.11** Dark absorption spectra for PYP in solution (blue) and on a crushed single crystal (red) measured by micro-spectrophotometer. Spectra are recorded from 25 accumulations of 25 $\mu$s exposures. The absorption maxima are at 446 nm and 449 nm in solution and crystal, respectively. For a crystalline sample, absorption at a laser wavelength of 485 nm is approximately ten times smaller than at 449 nm.

A typical absorption spectrum obtained from PYP in solution (2 mg/mL) and on a crushed crystal is displayed in Fig. 3.11. Compared to solution, the absorption maximum for a crushed crystal is shifted by 3 nm.

3.5.2.2 Time-resolved spectra on PYP solution

Fig. 3.12 shows results from a time-series of absorption spectra collected from PYP solution at 10 time delays from 30 $\mu$s to 2 s after reaction initiation. The dark spectra obtained before and after the time-series are shown in the inset of Fig. 3.12A, and it is
apparent that no permanent bleaching has occurred. For the early time points around 100 µs (colored black) a peak evolves at 485 nm and a sharp decrease is observed at 446 nm. As time progresses, the peak corresponding to 485 nm disappears and the 446 nm peak begins to recover. From the SVD analysis, two significant right singular vectors (rSVs) representing two kinetic processes are obtained as displayed in Fig. 3.12B. After global fitting of the rSVs by the sum of two exponentials, two relaxation times \( \tau_1 = 290 \mu s \) and \( \tau_2 = 0.38 \) s are determined. The first process corresponds to the pR to pB transition with a lifetime of 290 µs. The second process represents the recovery from the pB state to the ground state, pG.

### 3.5.2.3 Time-resolved spectra on PYP crushed single crystal

Difference spectra obtained from crushed single crystal at 0 ºC (Fig. 3.13A) and 30 ºC (Fig. 3.13B) are plotted as a function of wavelength from 400-600 nm. The dark spectra

**Figure 3.12** A. Time-resolved difference absorption spectra from 30 µs – 2 s for PYP in solution at 295 K. Black spectrum: earliest time point; Blue spectrum: last time point. The inset shows the dark spectra collected before (red) and after (blue) the full time-series. No permanent bleaching has occurred. B. Two significant rSVs obtained from the SVD analysis of the spectral changes from 400 nm to 510 nm: 1\(^{\text{st}}\) rSV (red circles); 2\(^{\text{nd}}\) rSV (blue circles); 3\(^{\text{rd}}\) rSV (green circles). Global fitting of the 1\(^{\text{st}}\), 2\(^{\text{nd}}\) and 3\(^{\text{rd}}\) rSV are represented by red, blue, and green curves, respectively, and relaxation times are shown as vertical lines.
obtained before and after the time-series are shown in the insets of Fig. 3.13A & B. It is apparent that permanent bleaching is either absent or very small. rSVs obtained from the SVD analysis at both temperature settings are shown in Fig 3.14 C-D. At 0 °C three significant rSVs equivalent to three kinetic processes are present (Fig. 3.13C). Process (1)

Figure 3.13 Time-resolved difference absorption spectra on crystalline slurry at different temperatures. The black horizontal line represents the difference spectrum corresponding to the dark (reference) spectrum. Black colored spectrum: 1st time point; blue colored spectrum: last time point A. T = 0 °C B. T = 30 °C. C-D. Significant rSVs from the SVD analysis of the difference spectra obtained for PYP crystalline slurry at different temperatures. 1st, 2nd and 3rd significant rSVs are represented by red, blue and green spheres, respectively. Solid curves denote the global data fitting by the sum of exponentials. Respective relaxation times are represented as vertical lines. Relaxation times (red) are obtained from crystallography (n.o. means not observed). C. At 0 °C three processes are identified and labeled with their respective relaxation times ( , & ) D. At 30 °C two processes with their respective relaxation times ( & ) are labeled.
corresponds to the pR to pB transition with relaxation time $\tau_1 = 2.1$ ms. The pB to pG transition is biphasic and characterized by both processes 2 and 3. The values of $\tau_2$ and $\tau_3$ are 67 ms and 800 ms, respectively. Only 2 significant rSVs are observed at 30 °C. These two rSVs are equivalent to two processes (Fig. 3.13D). The relaxation time corresponding to the 1st process is 2.2 ms. A minor contribution from the 2nd process with $\tau_2 = 32$ ms is observed.

**Table 3.11** Relaxation times ($\tau_i$) and corresponding macroscopic rate coefficients ($\Lambda_i$) for PYP in solution and for a crushed single crystal.

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Solution</th>
<th>Crushed single crystal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>295 K</td>
<td>273 K</td>
</tr>
<tr>
<td>Relaxation times $\tau_i$ (ms)</td>
<td>0.29</td>
<td>380</td>
</tr>
<tr>
<td>Macroscopic rate coefficients $\Lambda_i$ ($s^{-1}$)</td>
<td>3.4x$10^3$</td>
<td>2.6</td>
</tr>
</tbody>
</table>

**References**

4 Discussion

4.1 Catalase: structure and function

Our nitrosylated catalase (Fig. 3.1C) is of significant importance because it appears to be a good model system to mimic compound 1 ($\text{Fe-H}_2\text{O}_2$)(Eq. 1.2) that cannot either be recognized. The position of NO relative to the heme plane in our nitrosylated complex indicates the location of $\text{H}_2\text{O}_2$ bound in the actual catalytic reaction.

DEANO (Fig. 2.29 A) is a universal NO generator that can be used for all kinds of iron containing proteins. With the use of DEANO (Fig. 2.29A) and our simple glass apparatus (Fig. 2.29B), we can successfully produce 100 ~ 200 mmol/L of NO. The solubility of NO in water is ~2 mmol / L (Wilhelm et al., 1977). With this solubility, a 100 mmol/L of NO corresponds to a partial pressure of ~ 50 bars. It is crucial to start with such high concentrations of NO because the dissociation rate of NO from the Cat-NO is very high (Tab. 4.1). As soon as the crystals are removed from the NO containing solution, they quickly begin to lose their bound NO by diffusion. In addition, catalase crystals have to be immersed in a cryo buffer before freezing them in the cryostream (Sec. 2.5.2). This process requires at least a few seconds, in which the NO may further diffuse out of the crystals. By immersing our catalase crystals in the tightly closed cavity (Fig. 2.29B), we could literally pressurized the crystal with NO. Pressurization of these crystals with such high NO concentration at pressures much higher than 1 bar, enables us to detect a substantial occupation of NO in the crystal.

Another advantage of our method is that with the 100 mmol/L DEANO, the concentration of NO is 2 orders of magnitude higher than that of oxygen. Even if all the
oxygen in solution is consumed in reaction with NO, plenty of nitric oxide is still available to form Cat-NO complex. Handling of NO as well as performing crystallographic experiments under complicated anaerobic conditions is not needed.

Table 4.1 On and off rates as well as binding constants (K) for several protein-NO (or CO) complexes. Table from Purwar et al., 2011.

<table>
<thead>
<tr>
<th>Protein</th>
<th>$k_{on}$ (M$^{-1}$ s$^{-1}$)</th>
<th>$k_{off}$ (s$^{-1}$)</th>
<th>$K_{CatNO} = (k_{on}/k_{off})$ (M$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mb(Fe$^{2+}$)CO$^a$</td>
<td>5.1×10$^5$</td>
<td>1.9×10$^2$</td>
<td>2.7×10$^4$</td>
</tr>
<tr>
<td>Mb(Fe$^{2+}$)NO$^a$</td>
<td>1.7×10$^7$</td>
<td>1.2×10$^{-4}$</td>
<td>1.4×10$^{11}$</td>
</tr>
<tr>
<td>Mb(Fe$^{3+}$)NO$^a$</td>
<td>4.8×10$^4$</td>
<td>42</td>
<td>1.14×10$^7$</td>
</tr>
<tr>
<td>Lb(Fe$^{2+}$)NO$^a$</td>
<td>1.2×10$^8$</td>
<td>2.0×10$^{-5}$</td>
<td>5.9×10$^{12}$</td>
</tr>
<tr>
<td>NP1(Fe$^{4+}$)NO$^c$</td>
<td>1.5×10$^6$ (pH 5)</td>
<td>0.18 (pH 5)</td>
<td>8.3×10$^6$ (pH 5)</td>
</tr>
<tr>
<td>NP4(Fe$^{3+}$)NO$^c$</td>
<td>2.1×10$^6$ (pH 5)</td>
<td>0.11 (pH 5)</td>
<td>2.0×10$^7$ (pH 5)</td>
</tr>
<tr>
<td>Cat(Fe$^{3+}$)NO$^d$</td>
<td>1.3×10$^7$</td>
<td>1.5</td>
<td>8.7×10$^6$</td>
</tr>
</tbody>
</table>

$^a$ Ref. (Rohlfs et al., 1988); $^b$ Ref. (Laverman et al., 2001); $^c$ Ref. (Andersen et al., 2000); The nitrophorins adopt a closed and open conformation at pH 5 and pH 8, respectively; $^d$ This work (Purwar et al., 2011)

NO binding has been observed for both ferrous and ferric iron-containing heme proteins (Tab. 4.1). It binds extremely tightly to ferrous compounds such as Mb (Fe$^{2+}$) or Lb (Fe$^{2+}$) with small dissociation rate coefficients $k_{off}$ (Tab. 4.1). This leads to a larger binding constant ($K = k_{on}/k_{off}$) on the order of 10$^6$ or higher (Tab. 4.1). A kinetic analysis (Purwar et al., 2011) of the dissociation of NO from Cat-NO shows that the binding constant $K_{CatNO}$ for our catalase (Fe$^{3+}$) is 8.7 x 10$^6$ M$^{-1}$ (Tab. 4.1). Surprisingly, $K_{CatNO}$ is higher than those typically obtained for ferric-iron proteins such as Mb (Fe$^{3+}$) whose $K_{metMb-NO}$ is 1.14 x 10$^3$ M$^{-1}$ (Tab. 4.1). The reasons for these differences can be explained on the basis of our structural results (Fig. 3.1). In case of metMb, a water molecule is bound at the 6$^{th}$ coordination site of the heme (Laverman et al., 2001), whereas our
Crystallographic results for the Cat-5 complex show that this site is free (Fig. 3.1B). For the NO binding to metMb, this water molecule has to be displaced before a molecule of NO can enter to that site. In catalase, NO can easily occupy the free coordination site (Fig. 3.1B). This most likely accounts for the higher $k_{on}$ in catalase as compared to that of metMb (Tab. 4.1).

Our geometric calculations show that the NO is slightly bent away from the heme normal with an angle of $\sim 20^\circ$ (Tab. 3.1 & 3.3). This could be another reason for large $K_{\text{Cat-NO}}$. It is proposed that the relatively large bending of NO may be due to the spatial overlap of NO orbitals with the molecular orbitals of the iron/porphyrin system. This has been reported also for some other model systems (Richter-Addo et al., 2001). Depending on the amount of electron density provided to the iron, iron gains some ferrous character (Richter-Addo et al., 2001). Ferrous iron tends to bind NO more tightly. A higher degree of overlap can provide additional stability to the nitrosylated catalase. This may account for the smaller dissociation rate coefficient of our Cat-NO adduct as compared to the metMb-NO (Tab. 4.1).

It can be suspected that the relatively large bending angle in the Cat-NO complex is coming from the reduction of iron due to photoelectrons generated by X-rays. From our micro-spectrophotometric results, the photoreduction of iron for all 3 crystal forms are evident (Fig. 3.2A-F). Reduction of iron due to photoelectrons is also observed for Nitrophorin (Ding et al., 1999). Changes in the metal oxidation state may trigger major structural changes to the catalase, for example, the large bending geometry of NO. Since our experiments are performed at cryogenic temperatures (Sec. 2.5.2), only minute atomic rearrangements can take place (Kurinov & Harrison, 1995). From that, it can be
assumed that our catalase structures are in their original ferric form, even if iron is reduced from Fe(III) to Fe(II).

The previously proposed mechanisms (Fita & Rossmann, 1985, Alfonso-Prieto et al., 2007) for catalase show that the porphyrin-Fe (III) complex (Eq. 1.2-1.3) must have a tendency to easily donate electrons to facilitate the formation of compound 1. Our nitrosylated complex gains a partial ferrous character due to the spatial overlap as described above. The complex has an excess electron density and can easily give away electrons. We speculate that the ferrous characterization is enhanced, if the iron accepts some electron density from the negatively charged hydroxyl oxygen of the distal residue Tyr-357 (Tab. 3.1).

The structural changes in all three crystal forms of the catalase are minute (Tab. 3.1 & 3.4) and the iron out of plane distances are almost zero upon binding of NH$_3$ and NO (Tab. 3.1). The changes in the positions of active site residues are also negligible (Tab. 3.1). Similar results were reported for bacterial catalase upon ligand binding (Gouet et al., 1996). Catalase’s rigidity throughout the catalytic cycle is responsible for its role as an effective electron transfer protein. The structural changes remain as small as possible to optimize the speed of the reaction. These results differ for other heme proteins such as myoglobin (Schmidt et al., 2005b, Nienhaus et al., 2005). Once the ligand is removed heme itself becomes dome-shaped and the iron moves out of the heme plane by $\sim$ 0.3 Å (Schmidt et al., 2005a, Ihee et al., 2005). This triggers the large structural changes observed in this protein.

In our catalase structures, more than 1000 water molecules are identified (Sec. 3.1.3 & Tab. 3.2). These water molecules are located in many cavities (Tab. 3.2). The presence of
such large number of cavities provide a network to the heme and may facilitate the diffusion of $H_2O_2$. A similar environment containing many cavities was observed for cytochrome c oxidase (Tab. 3.2) whose function is also based on substrate diffusion and electron transfer.

### 4.1.1 Significance for other heme proteins involved in electron transfer

Our results can also be applied to other heme proteins such as cytochrome-c nitrite reductase (ccNIR, Sec. 4.2.1) (Darwin et al., 1993). This enzyme has some features that are similar to the catalase (Fig. 1.2). For example, it also has hemes in its active site (Einsle et al., 1999, Youngblut et al., 2011) and its catalytic reaction is based on the transfer of electrons (Cole, 1996). The number of electrons involved is 6 whereas 4 electrons are transferred in catalase (Sec. 1.4 & Eq. 1.2). When ccNIR catalyzes the reduction of nitrite to ammonia (Berks et al., 1995), several intermediates form and decay (Cole, 1996, Einsle et al., 2002, Kostera et al., 2008, Youngblut et al., 2011). The Cat-NO complex may mimic some of these crucial intermediates. This may provide an insight into the function of this important enzyme.

### 4.1.2 Summary

The three dimensional structure of Cat-NO adduct can be successfully obtained. This provides not only a detailed and deeper understanding of the reaction mechanism of the catalase but also an insight into other electron transferring proteins. These results also pave the way for easy crystallographic experiments with the NO in general and for investigations on other proteins which weakly bind NO.
4.2 **Time-resolved study on PYP**

The structures of our catalase adducts are static. They already provide information about how this protein works as a bio-catalyst. The aim is to push time-resolved experiments to perform functional studies on the catalase where we observed binding and dissociation of the substrate and product, respectively, in real time. The catalase reaction is irreversible, very fast and cannot be initiated conveniently by ultra-short light pulses. Therefore, it is tedious to follow this reaction using time-resolved crystallography on an appropriate time-scale. However, enzymatic reactions depend on temperature and in many cases also on pH. To explore the influence of these parameters on the kinetics, we employed a model system (PYP) where reaction can be repeatedly initiated in a convenient way. Results from these experiments are discussed in the following sections. Since repeated exposure to X-rays produce radiation damage, its effect on the photocycle is also explored and is discussed first. At the end, an outlook to new experiments which can be performed at the next generation of X-rays sources is presented.

4.2.1 **Effect of radiation dose**

Our results from dose experiments exhibit how the damaging effects of X-rays and laser pulses on PYP can be quantified. The uniqueness of our results is the determination of a kinetic dose limit in addition to the conventional dose limit, which is typically used to estimate the dose sensitivity of a crystal (Southworth-Davies *et al.*, 2007, Rajendran *et al.*, 2011, Kmetko *et al.*, 2011).

The relaxation times from SVD and posterior analyses are in agreement up to the kinetic dose limit, $D_{1/2}^K$ (Fig. 3.6A & B). After $D_{1/2}^K$, relaxation times from the posterior analysis
start to deviate. This shows that SVD can extract correct relaxation times even at higher X-ray doses because it is able to accurately deal with very small occupancies such as 3-5\% (Schmidt et al., 2003). On the other hand, sensitivity of the posterior analysis to noisy data causes the aforementioned deviations (Schmidt et al., 2012).

Up to $D_{1/2}^K$, 36 complete Laue data-sets can be collected from a PYP crystal of moderate size (170 x 170 x 700 $\mu$m$^3$) without disturbing the kinetic and structural analysis (Fig. 3.5). For a successful SVD analysis, 3 data sets per order of magnitude in time should be collected (Schmidt et al., 2003). Accordingly, a complete data set covering 12 orders of magnitude in time can be collected using only one single crystal. This means, with proper experimental settings, a comprehensive time-series from picoseconds to seconds can reliably be obtained (Schmidt et al., 2010).

As shown in previous studies, the dose limit $D_{1/2}$ at room temperature not only depends on the number of X-ray exposures (Tab. 4.2) used for the data collection but also on the employed dose rate (Kmetko et al., 2011, Rajendran et al., 2011, Southworth-Davies et al., 2007). The $D_{1/2}$ decreases when the dose rate is increased (Meents et al., 2010, Rajendran et al., 2011). The possible reason is that at a higher dose rate, the crystal heats up. Hydrogen may be produced and continues to accumulate in the crystal. In our case, the instantaneous dose rate ($\sim 2.3 \times 10^{13}$ Gy s$^{-1}$) is immense; however, the damaging effects are not significant. This is because we waited for 4s between two X-ray pulses so that the crystal can cool down (Moffat et al., 1992). This significantly reduces the average dose rate, which greatly affects the dose limit (Tab. 4.2). In our case, the dose limit is large even without using any radical scavengers. The addition of these
scavengers, such as ascorbate, to the crystals may increase this dose limit $D_{1/2}$ (Barker et al., 2009).

Our $D_{1/2}$ is one of the largest dose limits reported so far at room temperature (Kmetko et al., 2011, Rajendran et al., 2011, Southworth-Davies et al., 2007). However, it is more than an order of magnitude smaller than the Henderson-limit ($200 \times 10^5$ Gy) and the Owen-limit ($430 \times 10^5$ Gy). Both of these limits are obtained at cryogenic temperatures.

The $D_{1/2}$ values at cryogenic temperatures are relatively large, because the secondary damaging effects such as diffusion of radicals are strongly inhibited at lower temperatures (Owen et al., 2006).

**Table 4.2 Calculated value of dose limits and employed dose rates for different type of proteins.**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Dose limit ($D_{1/2}$) (Gy)</th>
<th>Dose rate (Gy s$^{-1}$)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysozyme</td>
<td>$16.3 \times 10^5$</td>
<td>10</td>
<td>(Southworth-Davies et al., 2007)</td>
</tr>
<tr>
<td>Insulin</td>
<td>$2.20 \times 10^5$</td>
<td>1430</td>
<td>(Rajendran et al., 2011)</td>
</tr>
<tr>
<td>PYP</td>
<td>$16.7 \times 10^5$</td>
<td>600</td>
<td>(<strong>This work</strong>) (Schmidt et al., 2012).</td>
</tr>
</tbody>
</table>

No site-specific damage is found at any of the amino acid residues of PYP (including Asp, Glu and those containing sulphur) from the inspection of the dark difference maps ($D_{i+1} - D_1$) (Schmidt et al., 2012). Interestingly, others (Kmetko et al., 2011) reported specific damage even at room temperature. The only effect observed in our difference map is that the noise level increases with the increasing number of time-series (Schmidt et al., 2012). The possible reason for the absence of these localized damages is that the damaged molecules quickly lose their structural integrity at the room temperature, whereas at cryogenic temperature molecules are more stable. As a result, at room temperature, even if some specific radiation damage occurs, its signature disappears as if
the entire molecule is removed (Schmidt et al., 2012). This generates some strain in a crystal which causes an increase in the mosaicity at higher X-ray doses as observed in our case.

Another interesting observation is that our corrected scattering intensities decay linearly (Fig. 3.3B). For the interpretation of dose data others proposed a first order model (Southworth-Davies et al., 2007) that would result in an exponential decay rather than the linear one. The reasons for these discrepancies are not known. Therefore, more experiments are necessary to propose a model that can be widely used to interpret the dose data from time-resolved crystallography.

4.2.1.1 Significance to other proteins

For PYP, the $D_{1/2}^K$ is almost half of the $D_{1/2}$ (Fig. 3.5B, Tab. 3.7). This relationship may hold for other proteins such as myoglobin (Schmidt et al., 2005b) and clam hemoglobin (Knapp et al., 2006), which were also investigated using long time-series from time-resolved crystallography. Results from these experiments allow us to setup a proper Laue experiment for dose-sensitive specimens such as the ccNIR (Youngblut et al., 2011) as mentioned already in Sec. 4.1.1. This sensitive protein has been investigated in collaboration with the Pacheco group. Several efforts were made to determine its structure using monochromatic X-rays but remained unsuccessful. Crystals of ccNIR obstinately resist freezing. In addition, these crystals quickly deteriorate in a monochromatic X-ray beam at ambient temperatures. Structure determination became feasible only when short polychromatic X-ray pulses were used. This suggests that short-
pulsed Laue experiments may be a preferred technique for proteins, which are highly sensitive to higher doses of X-ray radiation.

4.2.1.2 Summary

With the kinetic dose limit $D_{1/2}^K$, the number of kinetically meaningful data sets can be determined from a moderately sized single crystal of PYP. The crystal size used for our experiment is very well suited for time-resolved crystallographic experiments. This information can be used as a prerequisite to setup time-resolved crystallographic experiments for other proteins.

4.2.2 Effect of pH

In addition to the kinetic dose limit, we also explored how pH affects the structures and kinetics of PYP. The effect of pH has been observed spectroscopically, but a structural analysis has never been performed before.

At pH 4.0, the relaxation time for the $pB \rightarrow pG$ transition is increased by two orders of magnitude than that at neutral pH (Fig. 3.7B & 3.7E; Tab. 3.9). This is consistent with the previous studies which show that the pB relaxation slows down with acidification (Borucki et al., 2006, Genick et al., 1997a, Hoff et al., 1997). The reasons for this significant deceleration of the dark state recovery from the pB state at low pH were not known before. The presence of a water molecule corresponding to our feature $\beta$ (Fig. 3.8I.E & 3.8II.G) in the pB state allows us to explain this slower recovery. For the chromophore to relax from its pB state to the ground state, this water molecule has to be removed. The water molecule acts as a transient inhibitor. As soon as the pB state forms, the water enters in the chromophore pocket after 200 $\mu$s of the state formation. Since the
two events, formation of pB and the appearance of water are closely related, the water binding must be faster than pB formation. Therefore, water appears at the instant pB is formed as observed from the structure of the pB intermediate (Fig. 3.8II.E & 3.8II.G). The water molecule stays in the pocket and exits after ~1s.

The water molecule identified in the chromophore pocket may be a hydronium ion (H$_3$O$^+$) that neutralizes the negative charge on the Glu-46. Consequently, Glu-46 becomes protonated. The neutral electrostatic environment in the chromophore pocket may be one of the reasons why a large fraction of molecules relaxes directly from pR to pG (Fig. 3.7B & 3.7E).

The water molecule denoted by $\beta$ feature (Fig. 3.8II.E & 3.8II.G) forms hydrogen bonds with both Tyr-42 and Glu-46 (Sec. 3.4.3). This observation enables us to explain the results obtained from the previous Fourier Transform Infrared (FTIR) spectroscopy (Brudler et al., 2001, Shimizu et al., 2006). This study shows that the peak belonging to

![Figure 4.1](image-url)

**Figure 4.1** Structural differences of the PYP chromophore in the pG and blue shifted intermediate state pB. **A.** In pG, the chromophore is trans form and Glu-46 is protonated. An extended length of the conjugated system (red dashed lines) is possible. **B.** In the pB state, the proton has moved from the Glu-46 to the chromophore. Consequently, the chromophore becomes protonated. The effective length of the conjugated system is shortened and responsible for the blue shift of this intermediate.
the Glu-46 carboxyl group shifts from 1759 cm\(^{-1}\) to 1747 cm\(^{-1}\) when the pH of its surrounding environment is lowered. The presence of the water molecule (β feature; Fig. 3.8I.E & 3.8I.G) can potentially increase the hydrophilic environment near Glu-46 at low pH and may be responsible for the change in the Glu-46 IR-peak. Others (Meyer et al., 2003) also speculated that the hydrogen bond between a solvent molecule and Glu-46 is responsible for the spectral form of the PYP intermediate.

In addition to the feature β, our pB state shows the presence of another feature α outside the chromophore pocket at pH 7 and pH 4 (Fig. 3.8I.E & 3.8I.G, 3.8II.E & 3.8II.G). This feature was also observed by others (Ihee et al., 2005, Schmidt et al., 2004) and was defined as a double conformation of Arg-52. Our improved data quality enables us to modify this interpretation. We attribute the feature α as a water molecule that forms a hydrogen bond to the phenolate oxygen of the chromophore (Sec. 3.4.3). With this, one can explain the blue shift of the pB state (Xie et al., 1996). The presence of this hydrogen bond can affect the electron transition of the chromophore. The bond nature between phenolate oxygen and Glu-46 start to resemble a single bond (Fig. 4.1B). The effective length of the chromophore becomes shorter (compare Fig. 4.1A & B). This may be responsible for the blue shift. A similar observation has also been shown for a green fluorescent protein (GFP) chromophore, whose absorption maximum is blue-shifted in its protonated state (Kirchhofer et al., 2009). A previous study of the pCA chromophore analog suggests that hydrogen bonds are responsible for the additional ~60 nm shift when the protein environment is changed from aprotic to protic (Espagne et al., 2006). A decrease in wavelength of ~15 nm is observed for the pB state when protein is denatured.
from its native form (Lee et al., 2010, Xie et al., 2001). In this form, more hydrogen bonds may form with additional water molecules leading to the extra blue shift of ~15nm. At pH 7.0, relaxation times from the SVD analysis are in agreement with previous studies (Ihee et al., 2005, Schmidt et al., 2005a). Ihee and coworkers reported the presence of two pB phases at neutral pH. The second phase pB₂ is much less occupied as compared to the first phase pB (Ihee et al., 2005). Similarly, a faint kinetic phase corresponding to pB₂ at 71 ms is observed in our case (Fig. 3.7A). The reason for this weak phase is that the pB₁ and pB₂ are structurally very similar. Only one rSV is obtained which shows the presence of one phase coming from pB in general.

A comparison between relaxation times at pH 7.0 and 4.0 shows that the pR relaxes faster at low pH (Tab. 3.9). This causes an acceleration of the pR to pB transition. Similar behavior has also been observed before (Genick et al., 1997b). The relaxation time for the pB → pG transition at pH 4 is larger by two orders of magnitude than that at pH 7 (Tab. 3.9). This is consistent with previous studies which also show that the pB relaxation slows down with acidification (Borucki et al., 2006, Hoff et al., 1997, Genick et al., 1997b).

At basic pH (pH = 9), PYP stops cycling through the signaling state pB. From the pB state it reverts directly to ground state. The corresponding relaxation time obtained either from the SVD analysis (τ₁) or from the posterior analysis (τₚ₁), is on the order of ~2 ms (Fig. 3.7C & F; Tab. 3.9). This time scale of τ₁ or τₚ₁ is approximately one order of magnitude larger than that at pH 4 and 7. The higher pH prevents the detachment of the chromophore head from the hydrogen-bonding network. As a result, the pR state cannot relax into another state and the chromophore reisomerizes on a time scale of milliseconds
(Tab. 3.9) as also reported by others (Genick et al., 1997b). The cis–configuration after the isomerization can be further stabilized for an extended period of time only if another state is occupied. This happens at pH 7 and 4, where the pB state is largely occupied. At these pH’s, the pB state stabilizes the cis configuration of the chromophore for an extended period of time.

4.2.2.1 Summary

The variation of pH has a significant effect on the kinetics of the intermediate states in general and on the signaling state in particular. The pB state is absent at pH 9.0 whereas its life time extends at lower pH’s. We speculate that PYP may act as a pH sensing device in the host bacteria. The structural analysis of the reaction intermediates reveals the features that account for the blue or red shift observed spectroscopically.

4.2.3 Effect of temperature

From the dose and pH experiments, it is shown that our results are kinetically meaningful. We use this knowledge in exploring the effect of temperature on the kinetics of PYP. Results from our temperature dependent study shows that the functionality of PYP or proteins in general can be controlled by tuning the temperature. Not only structures of intermediates and chemical kinetics can be obtained but also thermodynamic properties can be extracted using the transition state equation (TSE, Eq. 4.1). The temperature dependent time-resolved study pushed crystallography to a new frontier. Results from both crystallographic and spectroscopic techniques are complementary to each other and can be tied together for a better understanding of PYP. Spectroscopic results also corroborate the findings from crystallography.
A comparison of relaxation times from time-resolved crystallography at -40 °C and 25 °C is shown in Fig. 3.10A. At higher temperature, the number of kinetic phases is reduced from 4 to 3. This is because at elevated temperatures (Appendix E), the photocycle of PYP speeds up. The fastest kinetic phase is shorter than our time-resolution and therefore cannot be detected. A decrease in the number of kinetic phases at temperatures higher than 50 °C is not due to a further acceleration of the photocycle (Schmidt et al., 2013). Protein starts to unfold and potentially more pB states are formed that superpose to a longer pB kinetic phase.

**Table 4.3 Rate of recovery from the signaling state pB to the dark state pG for PYP photocycle determined using time-resolved spectroscopy. The observed rates are compared with those reported by others.**

<table>
<thead>
<tr>
<th>Transition state</th>
<th>Solution</th>
<th>Crystal</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photostationary state → pG</td>
<td>(283 K) 666 ms&lt;sup&gt;a&lt;/sup&gt;</td>
<td>NA</td>
<td>(Meyer et al., 1989)</td>
</tr>
<tr>
<td>Photostationary state → pG</td>
<td>(283 K) 350 ms&lt;sup&gt;b&lt;/sup&gt;</td>
<td>(283 K) 454 ms&lt;sup&gt;d&lt;/sup&gt;</td>
<td>(Ng et al., 1995)</td>
</tr>
<tr>
<td>Photostationary state → pG</td>
<td>(283 K) 363 ms&lt;sup&gt;c&lt;/sup&gt;</td>
<td>(283 K) [191 ms &amp; 714 ms]&lt;sup&gt;e&lt;/sup&gt;</td>
<td>(Yeremenko et al., 2006)</td>
</tr>
<tr>
<td>pB2 → pG</td>
<td>(293K) 360 ms</td>
<td>(293K) ~20 ms</td>
<td>This work</td>
</tr>
<tr>
<td>pB1 or pB2 → pG</td>
<td>(295K) 380 ms</td>
<td>(273K) Biphasic [67 ms &amp; 800 ms] (303K) Biphasic [2.2 ms &amp; 32 ms]</td>
<td>This work</td>
</tr>
</tbody>
</table>

<sup>a</sup> 5 mmol/L sodium phosphate pH 7.0; <sup>b</sup> 10 mmol/L sodium phosphate pH 7.0; <sup>c</sup> 50 % saturated ammonium sulfate, 283K; <sup>d</sup> data fitting with one exponential; <sup>e</sup> data fitting with two exponentials; <sup>f</sup> from the global fitting of rSVs by a sum of exponentials, NA not applicable.

Our spectroscopic relaxation times for PYP in solution are compared with those reported by others (Tab. 4.3). The final decay (380 ms) obtained in our case is in very good agreement with the relaxation time (360 ms) observed by Yeremenko et al., (Tab. 4.3). This validates the design of our micro-spectrophotometer. In addition, our relaxation time for pB→pG transition also matches with that for photostationary to pG transition (Ng et
*al., 1995*); however, these results cannot be directly compared. This is because these relaxation times are obtained at different temperatures.

A comparison of spectroscopic relaxation times on a crystalline sample is not straightforward. Our results show that at low temperature, such as 273 K, the pB→pG transition is biphasic. Between the two kinetic phases, the 2nd one (800 ms) agrees with that (714 ms) reported previously (Tab. 4.3). The pB→pG transition at 303K is also biphasic where the 2nd phase (32ms) is consistent with a single kinetic phase obtained at 293K (Tab 4.3). In addition to that, a comparison of spectroscopic relaxation times between PYP in solution and on crystals provides an insight how the crystalline phase may affect the protein kinetics.

For the crystalline samples, two pB phases are observed spectroscopically at 0 °C whereas a single pB phase is detected from crystallography (compare Fig. 3.13 & Appendix E). This is because spectra are sensitive to electronic transitions whereas diffraction patterns are more sensitive to the structural changes. Consequently, even small electronic changes in the chromophore region can be detected from spectroscopy.

From the Arrhenius plot of macroscopic rate coefficients $\Lambda_i$ obtained from time-resolved crystallographic data (Fig. 3.10B-D), we determined the activation energies for the processes characterized by $\Lambda_2$ to $\Lambda_4$. The $\Lambda_i$ are linear combinations of the microscopic rate coefficients of the underlying chemical kinetic mechanism. For a precise calculation of thermodynamic variables, microscopic rate coefficients ($k_i$) must be determined. For this, a plausible kinetic mechanism that can properly interpret the data must be selected. Posterior analysis is applied for the selected mechanism. The mechanism used for the posterior analysis to our temperature dependent crystallographic data is shown in Fig.
3.9. The procedure to extract $k_i$ using this analysis is outside the scope of this thesis. This has been described in detail by a recent publication (Schmidt et al., 2013). Once $k_i$ are known for a proposed kinetic mechanism, their temperature dependence can be described using the transition state equation which is given by

$$k = \frac{RT}{N_A h} e^{\frac{\Delta S^\#}{R}} e^{\frac{-\Delta H^\#}{RT}},$$

where $R$ is the gas constant, $N_A$ is the Avogadro's number, $T$ is temperature, $h$ is the Planck's constant and $\Delta S^\#$ and $\Delta H^\#$ are the entropy and the enthalpy differences from an initial state to the transition state, respectively. From the plot of $k_i$ against temperature, thermodynamic variables such as enthalpy, entropy and Gibbs free energy differences to the transition state can be extracted (Schmidt et al., 2013). This is successfully applied to our time-resolved crystallographic data. Similarly, thermodynamic parameters can be extracted from time-resolved spectroscopy. However, more spectroscopic experiments performed at several temperature settings are necessary.

4.3 Future applications

As shown from our time-resolved crystallographic and spectroscopic experiments the time-resolved approaches are suitable for investigating cyclic reactions such as PYP, light sensitive complexes of heme proteins (Bourgeois et al., 2007, Schmidt et al., 2005b, Knapp et al., 2006, Srajer et al., 1996, Key et al., 2007) and photosynthetic reaction centers (Neutze & Moffat, 2012). Application of these techniques to enzymatic reactions is difficult because the reaction initiation on an ultra-short time scale is not straightforward. For that, a substrate has to be provided in an ultrafast manner. Once the reaction is over, the end products must be removed. Also, the new substrate has to be
loaded so that it can be freshly exposed to the enzyme. Another challenge for investigating the enzymatic reaction is that the reaction in an enzyme can be initiated only once due to the irreversibility of these reactions. To overcome these problems, either flow cells (Kurisu et al., 1997) can be used or a new crystal charged with substrate can be mounted each time. Both of these approaches are tedious and prevent rapid experiments.

Experiments on relatively faster time scales can be performed using caged substrates. In this method, an inactive caged substrate is soaked into the crystal so that the substrate can bind to the active site of the enzyme (Stoddard et al., 1998, Duke et al., 1994, Schlichting et al., 1990, Stoddard et al., 1991, Duke et al., 1992, Scheidig et al., 1992). A short and intense laser pulse is used to activate the caged substrate. The activation rate varies from nanoseconds to milliseconds (Adams & Tsien, 1993, Goeldner & Givens, 2005). The quantum yield of the activation may be low. A single laser shot may not activate a large fraction of the caged substrate. Chemical expertise is required to design and develop new caged substrates for various types of enzymes. These problems prevent the use of time-resolved crystallography as a routine tool to investigate these non-cyclic reactions.

Both challenges, fast reaction initiation and the requirement of multiple exposures during data collection, can be circumvented with the experiments performed at a modern facility using advanced techniques. New generation X-ray sources such as the XFEL (Fig. 4.2) deliver on the order of $10^{12}$ photons in a femtosecond X-ray pulse (Boutet et al., 2012). Having this flux, even a single pulse provides a diffraction pattern on a tiny crystal with sufficient signal to noise ratio. Such high flux enables us to perform time-resolved experiments on small crystals. In addition, the radiation damage is absent on fast femtosecond time scale (Chapman et al., 2011). For rapid reaction initiation, the easiest
way is to mix the enzyme with an active substrate and let the substrate diffuse into the crystal. To achieve the best time-resolution, the diffusion time ($\tau_D$) into the crystal must be as fast as possible. $\tau_D$ depends on the square of the crystal size (Schmidt, 2013). For example, $\tau_D$ is on the order of seconds for a crystal of size $\sim 300 \times 400 \times 500 \, \mu m^3$ and decreases to 1.5 $\mu$s for $0.1 \times 0.2 \times 0.3 \, \mu m^3$ sized crystals. Another requirement is fast mixing of the enzyme with the substrate. With newer mixing devices, time resolution can be as good as 150 $\mu$s (Cherepanov & de Vries, 2004). Injector technology is progressing rapidly to deliver enzyme-substrate mixture into the XFEL beam (DePonte et al., 2009, Park et al., 2006, Deponte et al., 2011, Gañán-Calvo et al., 2010, Sierra et al., 2012). Two events, mixing and injecting, must be carefully synchronized with X-ray pulses to probe changes in an enzymatic reaction.
This approach enables us to study enzymatic reactions that occur on faster time scales. One of the examples is catalase whose turnover time at ambient temperature is on the order of ~µs (Reid et al., 1981). In addition, this method can be applied to a wide range of the enzymes because catalytic rates for most of the enzymes are favorably in the range of few ms (Cornish-Bowden, 2012). The mix and inject technique paves a way to conveniently and routinely investigate enzymatic reactions on the faster (~µs) time scales.

References


5 Appendices

In this section, we provide the source code for the Andor basic program (Appendix A). This program was employed for the data acquisition as well as for initializing the iStar camera. During the data collection, timing is synchronized using two digital delay generators: DG545 and DG645. Only DG645 needed to be controlled by a program, therefore a Matlab graphical user interface (GUI) was created for this delay generator (Appendix B). In addition, Matlab programs were used to analyze the time-resolved spectroscopic data. The source codes for both of these programs are given in Appendices C & D. Finally, a figure from Schmidt et al., 2013 shows the relaxation times obtained from the time-resolved crystallography at different temperatures (Appendix E).
5.1 Appendix A: Program 1 for time-resolved spectroscopic data collection

This program is written in the Andor basic, a proprietary programming language. It is used to collect all the time-resolved absorption spectra at different time delays with minimal user input. All the desired time-delays are manually fed in this program using another Matlab script (not shown). Since this program can be initiated with an external trigger pulse, data acquisition can easily be synchronized with the Q-switch laser pulse.

```
RelaxationTime = 5 // Waiting time between two laser pulses
1 t = 1000; dim exp&[t]
2 gosub .createexposuretime //create the number of exposures from 2ms-3sec
3 exp&[1] = 0.002 ; i = 1
4 while(exp&[i] < 0.01) //exposure time from 2ms-9ms
5 exp&[i+1] = exp&[i]+0.001; i = i+1
6 wend
7 while (exp&[i] > = 0.01) && (exp&[i] < 0.1) // 10ms-90ms
8 exp&[i+1] = exp&[i]+0.01; i = i+1
9 wend
10 while (exp&[i] > = 0.1) && (exp&[i] < 1.0) // 100ms-1sec
11 exp&[i+1] = exp&[i]+0.1; i = i+1
12 wend
13 while (exp&[i] < 5.0) // 1sec-4sec
14 exp&[i+1] = exp&[i]+0.005; i = i+1
15 wend
16 num = i-1; print("Total number of exposure time: "; num)
17 d = 10; dim delay&[d] // 8 time delays in the range from 50 µs-1s
18 delay&[1]=0.00*10^000:delay&[2]=5.00*10^007:delay&[3]=2.06*10^008:delay&[4]=8.47*10^008:
20 delay&[9]=1.00*10^012:delay&[10]=0.00*10^000
21 dim rel_exp&[d] // Relevant exposure times(exposure times > (gate pulse width + pulse delay))
22 //Gate pulse width, a deciding factor for time-resolution
23 dim pw&[5]
25 170*10^6
26 // Select accumulation number based on gate pulse width
28 SaGain = 50; dim gain&[5] // Select the MCP gain, either same for each time delay or different
30 trigger = 2 //External triggering
31 setdatatype(6) //datatype is absorbance units; setacquisitionmode(2) //read in FVB mode
32 window = 100; i = 1; j = 1
33 for j = 1 to d
34 print(" next in loop j = ",j)
35 gosub .pulsewidthaccumulationifelseloop //Assign pulse width
36 while ( ((delay&[j] + pulsewidth)> = exp&[i]*10^12)&&(j<(d+1)))
37 i = i+1
```
wend
print(" i = ";i;" j = "j)
if ( (delay&[j] + pulsewidth) <= exp&[i]*10^12) then
rel_exp&[j] = exp&[i]
if(j = = d) then //*** For the dark after ***
rel_exp&[j] = exp&[1]; pulsewidth = pw&[1]
accumulationnumber = acmn&[1]; MCPgain = gain&[1]
endif
print(" For delay time (";j;") ",delay&[j]; ": exposure time(";i;") ");exp&[i]*10^12;":
relevant exposure time(";j;" );rel_exp&[j]*10^12
print(" pulsewidth = ";pulsewidth;" and accumulation number ";,accumulationnumber)
endif
next
TotalAccu = 0 // Initialization
gosub .takebackground
key(" TAKE REFERENCE NOW ") //wait for user input; gosub .takereference //take reference data
print("Total accumulation needed for Signal: ";TotalAccu);
MinutesSig = (TotalAccu*RelaxationTime)/60
print("Time required for Signal: ";MinutesSig/60;" hours")
key(" TAKE SIGNAL NOW ") // wait for user input; gosub .takesignal //take signal data
print ("Successfully DONE"); print(TotalAccumulation); saveoutput("output.txt")

// ******* Function.1 called in the main program
.pulsewidthaccumulationifelseloop
if (delay&[j] > 0) then
if (delay&[j] < 80*10^6) then
pulsewidth = pw&[1]; accumulationnumber = acmn&[1]; MCPgain = gain&[1]
print(" Time point ";j;": delay less than 80us, pulsewidth = ";pulsewidth)
endif
if ((delay&[j] > 80*10^6)&&(delay&[j] < 200*10^6))then
pulsewidth = pw&[2]; accumulationnumber = acmn&[2]; MCPgain = gain&[2]
print(" Time point ";j;": delay b/w 80us and 200us, pulsewidth = ";pulsewidth)
endif
if ((delay&[j] > 200*10^6) && (delay&[j] < 1*10^9))then
pulsewidth = pw&[3]; accumulationnumber = acmn&[3]; MCPgain = gain&[3]
print(" Time point ";j;": delay b/w 200us and 1ms, pulsewidth = ";pulsewidth)
endif
if ((delay&[j] > 1*10^9) && (delay&[j] < 500*10^9))then
pulsewidth = pw&[4]; accumulationnumber = acmn&[4]; MCPgain = gain&[4]
print(" Time point ";j;": delay b/w 1ms and 500ms, pulsewidth = ";pulsewidth)
endif
if (delay&[j] > 5*10^11)then
pulsewidth = pw&[5]; accumulationnumber = acmn&[5]; MCPgain = gain&[5]
print(" Time point ";j;": delay greater than 500ms, pulsewidth = ";pulsewidth)
endif
endif
return
// ******* Function.2 called in the main program
.takebackground ; print("Ready to take background")
for j = 1 to d
gosub .pulsewidthaccumulationifelseloop
65  print(" For delay time "; delay[j]; "; relevant exposure time "; rel_exp[j] * 10^12)
66  print(" pulsewidth = "; pulsewidth; "; accumulation number = "; accumulationnumber; " MCPgain
67   = "; MCPgain)
68  setDDG(trigger, 0, delay[j], pulsewidth, 0); setaccumulatenumber(accumulationnumber)
69  setGain(MCPgain); setexposuretime(rel_exp[j]); setacquisitiontype(1); run()
70  (window+j)_bg = #0_bg
71  if (j == 2) then
72     print(" Turn on the LASER for time series "); key(" Turn on the LASER for time series ")
73  endif
74  if (j == (d-1)) then
75     print(" Turn off the LASER and take dark "); key(" Turn off the LASER and take dark ")
76  endif
77  setaccumulatenumber(1); run()
78  (window+j)_sig = #0_sig; setaccumulatenumber(accumulationnumber); run()
79  (window+j)_sig = #0_sig; print(" Scan "; j; " is completed")
80  filename$ = "pyp_" + str$(window+j) + ".sif"; save(#(window+j), filename$)
81  filename2$ = "pyp_" + str$(window+j) + ".asc"; saveasciixy(#(window+j), filename2$, 2)
82  if (window+j)_sig = #0_sig; print(" Scan "; j; " is completed")
83  filename$ = "pyp_" + str$(window+j) + ".sif"; save(#(window+j), filename$)
84  filename2$ = "pyp_" + str$(window+j) + ".asc"; saveasciixy(#(window+j), filename2$, 2)
85  #0_sig = #0_sig; setaccumulatenumber(accumulationnumber); run()
86  TimeRem = TotalAccu * Relaxation Time/60; print(" Remaining time: "; TimeRem; " minutes")
87  next
88  save(#1000, "PYP_KS.sif"); copyxcal(#0, #1000)
89  saveasciixy(#1000, "test117.laser446.150uj.Wait4s.T285K.cov20.asc", 2); closewindow(#1000)
90  return
5.2 Appendix B: Graphical user interface for the timing synchronization

The graphical user interface is created in Matlab for the DG645 as mentioned. It is used to control the delay between flash lamp and Q-switch pulses. With this, one can also select the desired waiting time between two subsequent laser pulses.
5.3 Appendix C: Program 2 for time-resolved spectroscopic data analysis

The code performs the analysis of the time-resolved spectroscopic data. It requires an input file consisting of absorption spectra at different time-delays. Such file has already been produced from program.1. From these absorption spectra, difference spectra are first calculated and then brought to the same offsets. Spectral changes in a specified region are selected for the SVD analysis. Obtained right and left singular vectors are plotted against the time and wavelength axis, respectively.

```matlab
Clear all; clc;

% The given time range from 30us to 2s is distributed in 10 time points such that all of them are equally distributed on the logarithmic scale
tps=12; n1=log10(30)-6; n2=log10(2);%(1st & last timepoint:30us & 2 sec)
time=(logspace(n1,n2,tps-2)); (% Distribute in 10 time points)
tp(1)=0;tp(tps)=0; tp(2:tps-1)=time; t=tp';clear n1 n2 tp;

% Provide the ASCII file in which the first column is wavelength and remaining ones are the absorption spectra at certain time delays
Kinetic_Series_Xe='test161_Sol27.asc'
Detailedname='t161_RT.W5s.10tps25us~2s.Sol27.L446nm150uJ'
dark_figure=sprintf('Dark_%s.tif',Detailedname)
absorbance_figure=sprintf('Abs_%s.tif',Detailedname)
Difference_abs_figure=sprintf('DiffAbs_%s.tif',Detailedname)
LeftSingularValue=sprintf('Two_lSVs_%s.tif',Detailedname)
SingularValue=sprintf('SVs_%s.tif',Detailedname)
RightSingularValue=sprintf('Three_wrSVs_%s.tif',Detailedname)
delay=sprintf('Delay.tps_%s.mat',Detailedname);save(delay,'time');
wrSV=sprintf('wRSV_%s.mat',Detailedname)
diffSpec=sprintf('DiffSpectra_%s.asc',Detailedname)
FirRSVgrace=sprintf('firstrSVforGrace_%s.asc',Detailedname)

% Read the file & assigning variables
f=fopen(Kinetic_Series_Xe,'r') % read file
data=(fscanf(f,'%f',[tps+1 1024]))'; %Total Cols are 13 (1 lambda + 10 tps + 2 dark)
close(f);
PREwavelength=data(:,1); PREabsorbance=data(:,2:tps+1);
darkbefore=PREabsorbance(:,1);
darkafter=PREabsorbance(:,tps);

% Plot dark absorption spectra to examine the bleaching affects
figure(1);plot(PREwavelength,darkbefore,PREwavelength,darkafter)
title({'\bfDark spectra\rm'; sprintf('%s',Detailedname)},'fontsize',22,'fontname','Times New Roman')
xlabel({'\bfWavelength(\lambda)\rm','fontsize',22,'fontname','Times New Roman'})
ylabel({'\bfAbsorbance unit(A)\rm','fontsize',22,'fontname','Times New Roman'})
h=legend('dark before','dark after'); set(h,'fontsize',15,'fontname','Times New Roman')
set(gca,'fontsize',12,'fontname','Times New Roman')
```


Plot all the absorption spectra Vs wavelength

Calculate the difference absorption spectra and their offsets for each time delay

Obtain the absorption and difference spectra in the wavelength range from 400-500nm, where large spectral changes occur
Plot a time-trace of difference absorption spectra at a single wavelength such as 485 nm

```
DiffA_485nm = (diff(275,:))';
% The number 275 corresponds to 485 nm
Solution_22C = [time DiffA_485nm(2:tps-1)];
xlabel('Δt (s)','fontsize',22,'fontname','Times New Roman');
ylabel('ΔA (a.u.)','fontsize',22,'fontname','Times New Roman');
set(gca,'fontsize',16,'fontname','Times New Roman')
set(gca,'XLim',[10^-5 10],'XTick',logspace(-5,1,7))
save ('DiffA_485nm_Solu_22C.asc','Solution_22C', '-ASCII')
save ('DiffA_485nm_Solu_22C.mat','Solution_22C')
```

SVD calculation
```
[U,S,V]=svd(Diff,0);
figure(CCP4);plot(wavelength,U(:,1),'r-',wavelength,U(:,2),'b-');
title({'bfLeft singular vectors (U)';sprintf(' %s ',Detailedname)});
xlabel('Wavelength (λ)','fontsize',22,'fontname','Times New Roman');
ylabel('lSV','fontsize',22,'fontname','Times New Roman');
h=legend('1st lSV','2nd lSV');set(h,'fontsize',15,'fontname','Times New Roman');
set(gca,'fontsize',12,'fontname','Times New Roman')
saveas(figure(4),LeftSingularValue)
```

```
figure(5);plot(diag(S),'r-*');
h=legend('SVs');set(h,'fontsize',15,'fontname','Times New Roman');
title({'bfSingular values (S)';sprintf(' %s ',Detailedname)});
set(gca,'fontsize',12,'fontname','Times New Roman')
saveas(figure(5),SingularValue)
```

```
wrSV=V*S
figure(6);semilogx(time,wrSV(:,1),'r*',time,wrSV(:,2),'b*',time,wrSV(:,3),'g*');
title({'bfWeighted right singular vectors (V)';sprintf(' %s ',Detailedname)});
xlabel('delay time (s)','fontsize',22,'fontname','Times New Roman');
ylabel('S^2*rSV','fontsize',22,'fontname','Times New Roman');
h=legend('1st wrSV','2nd wrSV','3rd wrSV');set(h,'fontsize',15,'fontname','Times New Roman');
set(gca,'fontsize',12,'fontname','Times New Roman')
saveas(figure(6),RightSingularValue)
```

Save weighted S*rSVs for fitting routine Program.2
```
Three_wrSV=V*S(:,1:3);
wrSV=sprintf('S.RSV_%s.mat',Detailedname)
save (wrSV, 'Three_wrSV');
```

```
first_wrSV=Three_wrSV(:,1); firstrSVforGrace=[time first_wrSV]
save (FirRSVgrace, 'firstrSVforGrace', '-ASCII')
```
5.4 Appendix D: Program 3 for the fitting of right singular vectors using a sum of exponentials

This program globally fits the right singular vectors (rSVs) using a sum of exponentials. These rSVs were already calculated in the program 2 and used here as an input. As an example code, it is shown how two significant rSVs can be fitted using a sum of two exponentials.

```matlab
Clear all; clc;

load Delay.tps_t222.T30C.W5s.24tps25us~2s.C35.L446nm460uj.mat;
load S.RSV_t222.T30C.W5s.24tps25us~2s.C35.L446nm460uj.mat;

Detailedname =  't.222 30C'  % Formatted output
FittedData=sprintf( 'Fit_2RSV.2tau_%s.tif', Detailedname);
observed=sprintf( 'Obs_3RSV_%s.asc', Detailedname);
calculated=sprintf( 'Cal_2RSV.2tau_%s.asc ', Detailedname);

Assign the variables to the observed data
X=time;  Y1=Three_wrSV(:,1);  Y2=Three_wrSV(:,2); Y3=Three_wrSV(:,3);
figure(1);semilogx(X,Y1, '*r', X,Y2, '*b')
title(sprintf( 'Fitted Data %s ', Detailedname), 'fontsize', 25, 'fontname', 'Times New Roman')
xlabel('\Deltat (s)', 'fontsize', 22, 'fontname', 'Times New Roman');
ylabel('s*rSV', 'fontsize', 22, 'fontname', 'Times New Roman');
set(gca,'fontsize',16, 'fontname', 'Times New Roman')
set(gca,'XLim',[10^-5 10], 'XTick', logspace(-5,1,7), 'YLim',[-1.3 5])

The command calls a function fun_2RSV_2tau that starts at the point x0 and finds a minimum of the sum of squares of the functions. Calculate the new coefficients using LSQNONLIN.

[X0 = [-.0021 -.041 .1 .1 .1 -.1 .1 .1]'; % Initialization
Options = optimset('MaxIter',800,'MaxFunEvals',3000, 'TolFun',1e-10,'Algorithm','levenberg-marquardt','TolX',1e-10);
[x,resnorm,residual,exitflag,output]=lsqnonlin(@fun_2RSV_2tau,X0,[],[],options,X,Y1,Y2);

[abs(x(1)),abs(x(1)),abs(x(2)),resnorm])

h=text(v(1).*2,v(CCP4)*(0.4),RelaxationTimes); set(h,'fontname','Times New Roman','fontsize',12);

Generate a large number of time points such as 200 within in the same time range in which 10 points are originally generated. With this, smooth plot can be generated.

n1=log10(13)-6; n2=log10(CCP4);
SmoothTime=logspace(n1,n2,200); X=SmoothTime'; tau1=x(1);tau2=x(2); i=3;
A0=x(i);A1=x(i+1);A2=x(i+2);B0=x(i+3);B1=x(i+4);B2=x(i+5);
Y1calc=A0 + A1.*exp(X./tau1) + A2.*exp(X./tau2);
Y2calc=B0 + B1.*exp(X./tau1) + B2.*exp(X./tau2);
Hold on; semilogx(X,Y1calc,'r,X,Y2calc,'b');
h=legend('1st wrSV','2nd wrSV');
set(h,'fontsize',12,'fontname','Times New Roman');
saveas(figure(1),FittedData); hold off

observed_ascii=[time Three_wrSV]; save (observed,'observed_ascii','-ASCII')
Calculated_ascii=[X Y1calc Y2calc]; save (calculated,'Calculated_ascii','-ASCII')

****************************** THE FUNCTION fun_2RSV_2tau to be called ****************************
This function is called by lsqnonlin in the main body of the program. X, Y1 and Y2 are the input
variables that are passed to the lsqnonlin. x is the output vector.

function diff = fun_2RSV_2tau(x,X,Y1,Y2)
tau1=x(1);tau2=x(2); i=3; A0=x(i);A1=x(i+1);A2=x(i+2);B0=x(i+3);B1=x(i+4);B2=x(i+5);
diff = abs(A0 + A1.*exp(X./tau1) + A2.*exp(X./tau2) - Y1)+ abs(B0 + B1.*exp(X./tau1) +
B2.*exp(X./tau2) - Y2);
5.5 Appendix E: A figure showing the relaxation times obtained from time-resolved crystallography

The figure from Schmidt et al., 2013 shows the relaxation times obtained for PYP using time-resolved crystallography at different temperatures. The temperature range is from -40 °C to 50 °C. Upper panels show the rSVs from SVD analysis. From the global fit of these rSVs, the associated relaxation times are calculated. Lower panels show the concentration profiles obtained by kinetic modeling.
CURRICULUM VITAE

Namrta Purwar

Place of birth: Orai, U.P., India

Education

<table>
<thead>
<tr>
<th>Year</th>
<th>Degree</th>
<th>Institution</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006-Present</td>
<td>Ph.D. (Biophysics)</td>
<td>Department of Physics, UW-Milwaukee, USA</td>
<td>USA</td>
</tr>
<tr>
<td>2002-2004</td>
<td>M.S. (Physics)</td>
<td>Bundelkhand University, Jhansi, India</td>
<td>India</td>
</tr>
</tbody>
</table>

Dissertation Title

Structure and function of proteins investigated by crystallographic and spectroscopic time-resolved methods

Publications


7. Marius Schmidt; Vukica Srajer; Robert Henning; Hyotcherl Ihhee; Namrta Purwar; Jason Tenboer and Shailesh Tripathi (2013) Protein Energy Landscapes Determined by 5-Dimensional Crystallography. Acta Crystallographica Section D (Accepted).
Workshops attended and posters/Presentations

2013  Time-resolved experiments, Biophysics seminar, UW Milwaukee, WI

2012  Time-resolved spectroscopic study on photoactive yellow protein with newly designed micro-spectrophotometer, Biophysics seminar, UW Milwaukee, WI

2011  Structural and kinetic analysis of nitrosylated beef liver catalase & five dimensional crystallographic studies of photoactive yellow protein, Biophysics seminar, UW Milwaukee, WI

2011  Dynamic X-ray Scattering in Structural Biology held at BioCARS Advanced Photon Source, Argonne National Laboratory, Chicago, IL

2010  An X-ray Crystallographic Investigation of Nitric Oxide Binding to Beef Liver Catalase; Namrta Purwar, Jennifer McGarry, A. Andrew Pacheco, Marius Schmidt, American Crystallographic Association, Chicago IL

2010  Interaction of Nitric Oxide with beef liver catalase: Structural and Kinetic analysis, Biophysics seminar, UW Milwaukee, WI

2009  Seeing protein molecules change in real time using time resolved X-ray crystallography; Jennifer McGarry, Namrta Purwar, Marius Schmidt, A. Andy Pacheco, Undergraduate research symposium, UW Milwaukee, WI

2008  Time-resolved Macromolecular Crystallography held at BioCARS Advanced Photon Source, Argonne National Laboratory Chicago IL

Academic Awards/Achievements

2013  CGSA physics, David Lichtman memorial scholarship & UWM scholarship 1
2012  CGSA physics & David Lichtman memorial scholarship
2010-11  CGSA physics & UWM scholarship 1
2009  CGSA physics, Chancellor’s graduate student award & UWM scholarship 5ND
2008  CGSA physics
2006-07  Chancellor’s graduate student award
• Topper in B.Sc. in P.C.M. (Physics, Chemistry, Mathematics) in the B. B. College & Bundelkhand University, Jhansi (India)
• Second rank in M.Sc. (Physics) at B.B. College, Bundelkhand University, Jhansi
• Gold medal for highest marks in the Bundelkhand university, Jhansi